

Badanie efektywności psychoterapii metodą *Consumer Reports*: Analiza metodologii

Aleksandra Jaworska *

Instytut Psychologii, Uniwersytet im. Adama Mickiewicza, Poznań

A CONSUMER REPORT STUDY ON PSYCHOTHERAPY EFFECTIVENESS: ANALYSIS OF A METHODOLOGY

An analysis of the methodology used in the study on psychotherapy effectiveness conducted by *Consumer Reports (CR)* proved that the method had serious flaws. The main flaws undermining the credibility of the research findings are: biased sampling, lack of control groups and a quite imprecise and partial measurement tool. Each of these flaws taken separately is sufficient to consider the results of the *CR* study to be valueless because of low both internal and external validity. The methodology used in the study did not allow control of the factors limiting the two types of validity, such as: history, maturation, selection, statistical regression and mortality – for internal validity and non-representativeness of the sample and interaction of selection with an active factor (therapy) – for external validity. Because of this we can state that the main conclusion from the *CR* study: „psychotherapy is effective” is not justified as it lacks scientific grounds.

W 1996 r. *Czasopismo Psychologiczne* opublikowało tłumaczenie artykułu Martina Seligmana *Skuteczność psychoterapii. Badania ankietowe Consumer Reports* (Seligman, 1996b, oryginał ukazał się w numerze 50 *American Psychologist* z 1995 r). W artykule tym Seligman, który był głównym konsultantem badania, dokonał analizy metodologicznej zakrojonego na szeroką skalę badania skuteczności psychoterapii przeprowadzonego przez niezależną organizację konsumencką *Consumer Reports (CR)*, zajmującą się zakrojonymi na szeroką skalę badaniami produktów i usług, w których jednym z głównych źródeł informacji są opinie bezpośrednich użytkowników. W swym artykule Seligman wskazał szereg atutów przeprowadzonego badania, zasygnalizował jednak także szereg jego słabości metodologicznych. Mimo obecności tych ostatnich, konkluzja Seligmana była optymistyczna – badanie *CR* należy uznać za sukces, gdyż w wiarygodny sposób pozwoliło

ono potwierdzić skuteczność oddziaływań psychoterapeutycznych.

Jednakże wielu autorów wyraziło opinie zdecydowanie krytyczne wobec optymizmu Seligmana, poddające w wątpliwość wartość zastosowanej metodologii. Niniejszy artykuł stanowi przegląd opinii kwestionujących wartość metodologii chwalonej przez Seligmana oraz próbę krytycznej analizy spornych stanowisk.

W dyskusji nad metodologiczną wartością badania *CR* należy odwołać się do miejsca badań ankietowych opartych na opiniach pacjentów w typologii metod badawczych stosowanych w badaniach efektywności terapii oraz warunków stawianych poprawnym metodologicznie badaniom tego typu.

Dyskusja nad metodologią badania *CR* oparta była na powszechnym w literaturze amerykańskiej założeniu klasyfikującym badania nad efektywnością psychoterapii na dwa typy: *efficacy studies* i *effectiveness studies* (Seligman, 1995a). Próba znalezienia odpowiedników dla tych terminów w języku polskim jest dość trudna. Wypowiedzi Seligmana i innych autorów zajmujących się problematyką pomiaru skuteczności psychoterapii (Howard i in., 1996) wskazują, że *efficacy*

* Korespondencję dotyczącą artykułu można kierować na adres: Aleksandra Jaworska, Instytut Psychologii, Uniwersytet im. Adama Mickiewicza, ul. Szamarzewskiego 89, 60-568 Poznań.
e-mail: ale_jaw@smgkrc.pl

study jest odpowiednikiem eksperymentu laboratoryjnego, zaś *effectiveness study* – badania obserwacyjnego przeprowadzonego w warunkach naturalnych (badania terenowego), w których badacz nie ma możliwości manipulacji zmiennymi niezależnymi. Nathan, Stuart i Dolan (2000) w swym przeglądzie badań nad psychoterapią w ujęciu historycznym wskazują, że przez wiele lat większość badaczy koncentrowała się głównie na badaniach typu eksperymentalnego, przez wielu uznawanych za jedyne wiarygodne źródło naukowych danych odnośnie skuteczności psychoterapii. Jednak na początku lat 80-tych rozbieżności między sytuacją laboratoryjną a sytuacją rzeczywistą i związane z nimi ograniczenia związane z aplikacją wyników badań eksperymentalnych do rzeczywistości zaczęły być coraz wyraźniej zauważane, różnice między tymi dwoma typami metodologii były coraz bardziej podkreślane. Barlow (1996) w sposób zwięzły ujął definicje badań eksperymentalnych i terenowych nad efektywnością psychoterapii. Jego zdaniem, badania eksperymentalne nad psychoterapią dotyczą „rezultatów systematycznej oceny przeprowadzanej interwencji w kontrolowanej sytuacji badania klinicznego. Wnioski z takich badań podkreślają zwykle kwestie związane z trafnością wewnętrzną” (s. 1051). Z kolei badania kliniczne efektywności psychoterapii mają na celu „ustalić stopień możliwej generalizacji interwencji terapeutycznych o określonym poziomie skuteczności” na rzeczywiste warunki (s. 1055). Podobnego zdania jest Wampold (1997), który uważa, że badania typu *efficacy* służą do wykrycia związku przyczynowo-skutkowego w warunkach eksperymentalnych, natomiast badania typu *effectiveness* próbują określić, czy pacjenci odnoszą korzyści z psychoterapii przeprowadzanej w warunkach naturalnych. Badania skuteczności psychoterapii kładą zatem nacisk na trafność wewnętrzną oraz replikowalność, zaś badania efektywności – na trafność zewnętrzną oraz możliwość generalizacji wyników.

Zdecydowana większość przeprowadzonych badań nad efektywnością psychoterapii, których najbardziej znane przeglądy i meta-analizy zawarte są w pracach Lipseya i Wilsona (1993) oraz Smitha i Glassa (1977) to eksperymenty laboratoryjne. Na tego typu badaniach opiera się w znacznej mierze aktualna wiedza na temat skuteczności psychoterapii. Jednak zwolennicy metod terenowych, z Seligmanem na czele, podkreślając wskazane wyżej ograniczenia metod laboratoryjnych, dowodzą, że nie można opierać się tylko na badaniach typu *efficacy* – obraz problemu nie jest pełen, jeżeli nie bierze się pod uwagę, jak psychoterapia działa w warunkach naturalnych. Badanie przeprowadzone przez *Consumer Reports* uważane jest za najbardziej znany przykład badania typu terenowego.

PRZEBIEG BADANIA ORAZ JEGO NAJWAŻNIEJSZE WYNIKI

Badanie ankietowe *Consumer Reports* zostało przeprowadzone w USA na przełomie lat 1994 i 1995. Kwestionariusz zawierający pytania dotyczące efektywności leczenia problemów natury psychicznej przy pomocy psychoterapii oraz leków (oraz moduł z pytaniami dotyczącymi samochodów) został wysłany do około 180000 osób. O odpowiedź proszone były osoby, które w ciągu ostatnich trzech lat „doświadczyły stresu lub innych problemów emocjonalnych i w związku z tym szukały pomocy u przyjaciół, krewnych lub księży, a także profesjonalistów do spraw zdrowia psychicznego – psychologa, psychiatry, lekarza ogólnego lub grupy wsparcia”. Odsetek respondentów, którzy odesłali ankietę wyniósł 13% (ok. 22000 osób), z czego na pytania interesującej nas części kwestionariusza odpowiedziało ok. 7000 osób, czyli niecałe 4% wszystkich osób, które otrzymały ankietę. Spośród tych respondentów, 2900 osób (1.6% spośród osób, do których wysłano kwestionariusz) skorzystało z pomocy profesjonalistów w zakresie zdrowia psychicznego – 37% osób brało udział w terapii prowadzonej przez psychologa, 22% – przez psychiatrę, 14% – przez pracownika socjalnego, 9% – doradcę do spraw małżeńskich, zaś 18% było leczonych przez innego typu specjalistów (Seligman, 1995a). Kwestionariusz użyty w badaniu składał się z głównego modułu 26 pytań dotyczących pomocy uzyskanej od profesjonalistów w zakresie zdrowia psychicznego oraz podobnych, lecz mniej szczegółowych pytań dotyczących lekarzy ogólnych, leczenia farmakologicznego oraz grup samopomocowych¹.

Na podstawie danych otrzymanych w badaniu *CR* jego autorzy doszli do następujących wniosków (Seligman, 1995a, s.968-969, Kotkin, Daviet i Gurin, 1996, s. 1080):

- terapia prowadzona przez psychoterapeutów specjalistów była najczęściej skuteczna;
- im dłużej trwała terapia, tym lepsze były jej rezultaty;
- sama psychoterapia była równie skuteczna, jak psychoterapia połączona z leczeniem farmakologicznym;
- wszyscy specjaliści do spraw zdrowia psychicznego pomogli swoim pacjentom, przy czym doradcy do spraw małżeńskich radzili sobie nieco gorzej;
- bardzo skuteczne było leczenie w grupach Anonimowych Alkoholików;

¹ Ani *Consumer Reports*, ani *Consumer Union* nie chciały udostępnić autorce artykułu pełnej wersji oryginalnego kwestionariusza użytego w badaniu. Zatem wszystkie uwagi odnośnie budowy kwestionariusza i sposobu formułowania pytań opierają się na wzmiankach i odniesieniach zawartych w cytowanych publikacjach.

- żaden z rodzajów terapii nie okazał się lepszy niż inne w leczeniu określonej kategorii zaburzeń;
- lekarze rodzinni byli równie skuteczni jako profesjonalści w zakresie zdrowia psychicznego w terapii krótkoterminowej, natomiast ci ostatni byli bardziej skuteczni w terapii długoterminowej.

KONTROWERSYJNE KWESTIE METODOLOGICZNE W BADANIU CONSUMER REPORTS

Systematyczna analiza metodologicznej wartości badania metodą *Consumer Reports* składa się z omówienia czynników mających wpływ na trafność wewnętrzną i zewnętrzną badania. Analizie zostały poddane następujące kwestie:

- kto został poddany badaniu?
- jak zostało dokonane badanie?
- co uzyskano w rezultacie badania?

DOBÓR OSÓB BADANYCH

Kluczową kwestią w dyskusji nad wartością metodologii badania *CR* był problem doboru osób badanych, związany z reprezentatywnością próby i możliwością generalizacji otrzymanych w badaniu wyników na populację.

Seligman oraz analitycy *Consumer Reports* (Kotkin, Daviet i Gurin, 1996, Kotkin i Daviet, 1999) wielokrotnie podkreślali prawidłową metodologię doboru próby do badania. Autorzy ci twierdzą, że choć próba nie jest reprezentatywna dla całej populacji USA, można przypuszczać, że jest ona „reprezentatywna dla wykształconej klasy średniej, z której pochodzi większość osób poddających się psychoterapii” (Seligman, 1995a, s. 969). Prawidłowość doboru próby ich zdaniem potwierdzają:

- 1 – podobny jak w przypadku innych badań typu *CR* odsetek osób odpowiadających na pytania kwestionariusza (13%)
- 2 – ankieta walidacyjna, mająca na celu kontrolę stronnictwa doboru próby wynikającego z faktu, że część osób nie odesłała wypełnionego kwestionariusza – polegała ona na wysłaniu do pewnej grupy osób kwestionariusza z opłaconym znaczkiem pocztowym (podczas gdy w całej próbie koszt przesłania kwestionariusza spoczywa na osobie badanej). W ramach ankiety walidacyjnej wypełniony kwestionariusz przesłało 38% badanych, których odpowiedzi nie różniły się w istotny statystycznie sposób od odpowiedzi osób z próby głównej (Seligman, 1995a, s. 971).
- 3 – Zgodność odpowiednich odsetek respondentów z danymi epidemiologicznymi – spośród 13% osób, które odesłały wypełniony kwestionariusz, około 1/3

twierdziła, że miała problemy natury emocjonalnej/psychicznej. Dane te są zgodne z danymi zakrojonych na szeroką skalę badań epidemiologicznych Epidemiological Catchment Area z 1984 r. oraz National Comorbidity Study z 1994 r., według których w ciągu roku około 30% dorosłej populacji USA dotknięte jest pewnego rodzaju zaburzeniami natury psychicznej. Z kolei wysokości odsetka osób, które odpowiedziały na pytania kwestionariusza (ok. 4% z całej próby) jako osoby, które w ciągu ostatnich 3 lat skorzystały z pomocy specjalisty w zakresie zdrowia psychicznego zgodny jest z danymi uzyskanymi w badaniu National Medical Care Utilization and Expenditure Survey z roku 1984, przeprowadzonego na próbie 6000 gospodarstw domowych w USA, które wykazało, że 4% Amerykanów korzystało z tego typu pomocy w r.1980 (Seligman, 1996a, s. 1073).

Ubiegając swoich krytyków, Seligman sam wskazał kilka potencjalnych źródeł błędów stronnictwa próby (1995a):

- Wyniki osób poddawanych leczeniu w grupach AA, których odpowiedzi dowodziły niezwykle wysokiej skuteczności tej formy terapii mogą być rezultatem z jednej strony faktu, że osoby leczące się w grupach AA zwykle zachęcane są do członkostwa trwającego całe życie, w związku z tym istnieje większe prawdopodobieństwo, że osoby pytane o pomoc, z której skorzystały w ciągu ostatnich 3 lat, są osobami, w stosunku do których ten typ leczenia okazał się skuteczny.

- Należy także dopuścić możliwość, że u alkoholików, którym terapia w grupie AA nie pomogła, mogły pojawić się poważne zaburzenia funkcjonowania, które miały wpływ na fakt wzięcia przez nich udziału w badaniu – istnieje prawdopodobieństwo, że próba badania zawiera nadreprezentację osób, w stosunku do których leczenie alkoholizmu okazało się skuteczne, zaś osoby, wobec których leczenie to okazało się nieskuteczne są niedoreprezentowane.

- Kolejnym możliwym źródłem stronnictwa próby, mającym wpływ na wnioski odnośnie większej skuteczności terapii długoterminowej, jest prawdopodobieństwo, że osoby, którym terapia planowana jako długoterminowa nie pomaga, przerywają leczenie po kilku miesiącach. Istnieje zatem możliwość, że uczestnikami terapii długoterminowej są te osoby, którym terapia ta pomagała i które w związku z tym miały motywację do jej kontynuacji.

- Innym ograniczeniem reprezentatywności próby wskazanym przez Seligmana jest fakt, że cała próba składała się z osób, które dokonały wyboru rodzaju leczenia, a zatem zdawały sobie sprawę z istnienia problemu (problemów) oraz prawdopodobnie wierzyły w skuteczność podjętej terapii. Osoby te, to osoby raczej aktywne, ukierunkowane na rozwiązywanie problemów (problem – solvers). Fakt ten ogranicza możliwość generalizacji wyników badania na osoby,

które zaprzeczają, by miały jakieś problemy natury emocjonalnej/psychicznej oraz które nie dokonują wyboru typu leczenia oraz terapeutę.

Wszystkie te słabe punkty doboru próby zdaniem Seligmana nie podważają jednak wiarygodności wyników badania i nie jest prawdopodobne, by były źródłem poważniejszych artefaktów (Seligman, 1995a, Seligman, 1996a, s. 1073).

Metodologia doboru próby została skrytykowana przez wielu autorów (Jackobson i Christensen, 1996, Brock, Green, Reich i Evans, 1996), którzy zarzucali próbie przede wszystkim brak reprezentatywności. Jackobson i Christensen wskazują, że nie można odrzucić hipotezy, że osoby, którym psychoterapia pomogła były bardziej skłonne do wzięcia udziału w badaniu i udzieleniu odpowiedzi na pytania kwestionariusza *CR*. Autorzy nawiązują tym samym do problemu selektywności próby jako źródła błędu jej stronniczości. Podobnie twierdzą Brock, Green i Reich (1998) – dowodzą oni, że fakt, iż dane stanowiące podstawę do wysnucia wniosków z badania pochodzą od bardzo małego ułamka respondentów, którzy wzięli udział w badaniu, zaledwie 4%, uprawdopodobnia podejrzenia, że wśród 96% pozostałych osób, które nie udzieliły odpowiedzi na pytania z interesującej nas części kwestionariusza było wielu takich respondentów, którzy prawdopodobnie korzystali z jakiejś formy terapii, która zakończyła się niepowodzeniem lub która była w niewielkim stopniu skuteczna, a którzy z różnych powodów nie wypełnili długiego kwestionariusza i nie odesłali go z powrotem do *CR*.

Choć Seligman sygnalizował problem stronniczości próby spowodowanego jej możliwą selektywnością w odniesieniu do osób leczących się w grupach AA, odnośnie pozostałych biorących w badaniu osób wypowiedział się w tym względzie jedynie ogólnikowo, dopuszczając możliwość wystąpienia tego typu błędu wśród pozostałych respondentów, jednak „w mniejszym stopniu” (Seligman, 1995a, s. 971). Twierdzenie to nie zostało przez Seligmana uzasadnione, a przecież wzbudza duże wątpliwości – fakt nadreprezentacji w próbie osób, odnośnie których psychoterapia okazała się skuteczna, ma kluczowe znaczenie dla wartości metodologicznego badania, a zwłaszcza jej najważniejszego wniosku dotyczącego skuteczności psychoterapii w ogóle. Kotkin, Daviet i Gurin (1996, s. 1081) bronili stanowiska głoszącego, że fakt braku odzewu ze strony znikomej części osób, których zaburzenia są tak poważne, by uniemożliwić im wzięcie udziału w badaniu, nie ma istotnego wpływu na reprezentatywność próby, skoro podkreślają oni wyraźnie, że wyniki badania nie odnoszą się do tej niewielkiej części populacji.

Na kwestię reprezentatywności próby można także spojrzeć z szerszej perspektywy, która pozwala uniknąć wdawania się w spekulacje dotyczące głębokości ewentualnych zaburzeń osób, które nie odpowiedziały na

ankietę *CR*. W doborze próby w badaniu *CR* ujawnił się bowiem w całej rozciągłości problem ochotników – osób dobrowolnie biorących udział w badaniu. Stronniczość próby spowodowana udziałem w badaniu ochotników jest poważnym czynnikiem rzutującym na możliwość generalizacji uzyskanych wyników (Brzeziński, 1996). Jak wykazały liczne badania empiryczne, próba składająca się z ochotników składa się z osób charakteryzujących się szczególnymi cechami – są to zwykle osoby o wyższym poziomie wykształcenia, wyższym statusie społecznym, wyższym poziomie inteligencji, wyższym poziomie zmiennej aprobaty społecznej oraz bardziej zsocjalizowane (Rosenthal, Rosnow, 1991). Z kolei badania (cytowane przez Mangione, 1999, s. 82) nad błędem braku odpowiedzi pozwalającym się w badaniu metodą ankiety pocztowej pozwoliły ustalić, że osobami, które nie odpowiadają na nadesłane im kwestionariusze, są zazwyczaj osoby słabo wykształcone, starsze, stanu wolnego, płci męskiej, a także „osoby, którym wydaje się, że nie pasują do tematyki, jakiej dotyczy ankieta (na przykład abstynenci proszeni o wypełnienie ankiety o alkoholizmie [...])”. Możemy zatem przypuszczać, że tego typu osoby były mniej skłonne do udziału w badaniu. Choć problem różnic między tymi dwiema grupami osób miała rozwiązać ankieta walidacyjna, należy pamiętać, że respondentami w tej części badania byli wszakże także ochotnicy.

Rozpatrując problem reprezentatywności próby należy pamiętać, że im większe są różnice między osobami, które wzięły udział w badaniu oraz tymi, które spełniały warunki uprawniające je do udziału w badaniu (problemy emocjonalne oraz skorzystanie z pomocy terapeutycznej w ciągu ostatnich 3 lat), a które w nim udziału nie wzięły, tym większe prawdopodobieństwo, że wnioski z badania odnosić się będą jedynie do wąskiej grupy osób, których charakterystyka nie tylko demograficzna, ale i psychologiczna odpowiada profilowi osób, które odpowiedziały na ankietę.

Nie można wykluczyć, że rozkład zmiennej niezależnej w postaci typów i głębokości zaburzeń emocjonalnych poddawanych terapii, na które cierpieli ochotnicy biorący udział w badaniu różnił się od rozkładu tej zmiennej w populacji – zmienna ta nie była poddana kontroli. Warto tu nadmienić, że inaczej niż w badaniach laboratoryjnych, do których w celu zwiększenia możliwości kontroli zmiennych ubocznych wybiera się osoby, u których występuje tylko jeden typ zaburzeń, o ściśle opisanych objawach, w badaniu *CR* brały udział osoby, u których wystąpił jednocześnie więcej niż jeden problem natury emocjonalnej/psychicznej (co Seligman podkreślał jako jedną z jego głównych zalet badania, zgodną z jego naturalistycznym charakterem).

Jednakże brak kontroli tak istotnej zmiennej, jak typ i głębokość poddanego terapii zaburzenia (zaburzeń) emocjonalnego (emocjonalnych) pozwala na wysunięcie hipotezy, że psychoterapia podziałała tylko

w przypadku niektórych typów problemów natury psychicznej (np. tych mniej poważnych) i to właśnie osoby z określonym typem zaburzeń wzięły udział w badaniu.

Podsumowując, można zatem wskazać szereg czynników, które uniemożliwiają generalizację wyników badania CR na populację osób cierpiących na zaburzenia emocjonalne, które brały udział w psychoterapii:

- ogólne ograniczenia płynące ze szczególnych cech osób ochotniczo biorących udział w badaniu, a także hipoteza, że przynajmniej niektóre z tych cech (np. wyższej samoświadomości własnych problemów oraz wyższy poziom motywacji do rozwiązywania swoich problemów) są dodatnio skorelowane z opisywanymi przez nie pozytywnymi efektami psychoterapii, a zatem stanowią niekontrolowane istotne zmienne niezależne;

- nieznany i niekontrolowany poziom tak istotnej zmiennej niezależnej, jak typ i głębokość zaburzeń osób badanych poddawanych terapii. Jak sugerują Jakobson i Christensen (1996, s. 1034-1035) poziom zaburzeń u tych osób mógł być dość nieznaczny, w związku z czym były one znacznie bardziej podatne na odniesienie korzyści z leczenia lub na spontaniczną remisję. Twierdzą oni, że uzasadnione jest przypuszczenie, że stan wielu spośród badanych, którzy relacjonowali pozytywną zmianę, poprawiłby się samoistnie bez udziału w żadnej terapii.

Nie można zatem wykluczyć hipotezy o zwiększonym prawdopodobieństwie tego, że udział w badaniu wzięły osoby, którym psychoterapia pomogła oraz posiadające cechy sprzyjające wystąpieniu pozytywnych efektów terapii. Tak liczny zbiór warunków ograniczających generalizację wyników badania na populacji osób biorących udział w psychoterapii sprawia, że ogólny wniosek z badania, iż „psychoterapia jest skuteczna” nie jest uprawniony.

KWESTIA GRUP KONTROLNYCH ORAZ RANDOMIZACJI

Drugim istotnym problemem szeroko omawianym w dyskusji nad metodologią badania CR była kwestia grup kontrolnych. Wiąże się on z odpowiedzią na dwa pytania, a mianowicie: czy grupa kontrolna jest konieczna, aby wyciągnąć uprawnione wnioski na temat skuteczności leczenia oraz: jaka powinna być natura grupy kontrolnej (Nathan, Stuart, Dolan, 2000, s. 965).

Kwestią bezsporną jest, że obecność grupy kontrolnej (porównawczej) jest niezbędna dla zapewnienia trafności wewnętrznej badania typu eksperymentalnego czy *quasi*-eksperymentalnego. Plany *quasi*-eksperymentalne, nie uwzględniające obecności grupy kontrolnej, narażone są na poważne zagrożenia dla trafności wewnętrznej badania, takich jak oddziaływanie czynnika historii (problem spontanicznej remisji), dojrzewania lub regresji statystycznej (Cook i Campbell, 1979).

W badaniach eksperymentalnych i *quasi*-eksperymentalnych nad efektywnością psychoterapii powszechnie stosuje się kilka typów grup kontrolnych, do których należy: grupa placebo (w której pacjenci poddani się oddziaływaniu, co do którego przekonani są, że ma wartość terapeutyczną, podczas gdy w rzeczywistości takowej nie posiada), grupa osób oczekujących na terapię (*waiting list*), bądź też grupa (grupy) osób leczonych standardową metodą terapeutyczną o sprawdzonej skuteczności (Nathan, Stuart i Dolan, 2000).

W badaniu *Consumer Reports* nie zastosowano żadnego rodzaju kontroli w postaci grupy kontrolnej czy to osób, które cierpiały na zaburzenia emocjonalne i nie podjęły leczenia, ani nie korzystały też z pomocy nieprofesjonalnej czy też osób, które nie zwróciły się o pomoc do psychologa czy psychiatry, lecz korzystały z pomocy znajomych czy osób profesjonalnie zajmujących się tego typu problemami.

Seligman (1995a) przyznał otwarcie, że istnieją pewne problemy natury metodologicznej badania CR wynikające z braku grupy kontrolnej – brak ten nie pozwala np. stwierdzić bezpośrednio, czy istnieje różnica między wynikami leczenia przez profesjonalistę w zakresie zdrowia psychicznego a wynikami „leczenia” polegającego na rozmowie z współczującym przyjacielem, lub poprawą wynikającą z samego tylko upływu czasu. W badaniu CR osoby, które zwróciły się o pomoc do nieprofesjonalistów – znajomych lub duchownych – nie były, niestety, proszone o odpowiedzenie na pytania dotyczące rezultatów tego typu pomocy.

Jednak zdaniem Seligmana, fakt braku grupy kontrolnej nie jest istotnym brakiem metodologicznym ważącym na wiarygodności wyników badania, jako że zawierało ono pewne elementy mogące spełniać rolę wewnętrznych grup kontrolnych. Seligman wskazuje cztery sposoby kontroli, które mogą jego zdaniem z powodzeniem rewalidować brak zewnętrznej grupy kontrolnej:

- Rolę takiej grupy mogą np. stanowić doradcy do spraw małżeńskich, którzy okazali się znacznie mniej skuteczni w leczeniu wszystkich rodzajów zaburzeń niż psychologowie, psychiatrzy czy pracownicy socjalni. W relacjach badanych z doradcami małżeńskimi pojawiały się czynniki niespecyficzne, mogące mieć wpływ na skuteczność leczenia, takie jak związek pacjenta z terapeutą, uwaga skierowana na pacjenta oraz upływ czasu.

- Drugim elementem kontrolnym jest, zdaniem Seligmana, krzywa reakcji dowodząca, że terapia jest tym skuteczniejsza, im jest dłuższa.

- Trzecim elementem kontrolnym są dane dowodzące, że sama psychoterapia jest równie skuteczna w leczeniu wszelkiego rodzaju zaburzeń, jak psychoterapia w połączeniu z leczeniem farmakologicznym. Seligman uważa, że w takiej sytuacji wprowadzenie grup kontrolnych nie jest już potrzebne, jako że przeprowadzono

już tyle badań porównujących skuteczność leków i placebo, z których jasno wynika, że farmakoterapia jest skuteczna, że „uzasadnione staje się przypuszczenie, że jedynie raz jeszcze okazałoby się, że psychoterapia jest skuteczniejsza niż placebo” (*ibidem*).

– Czwartego typu grupę porównawczą mogą, zdaniem Seligmana, stanowić doradcy rodzinni, którzy okazali się mniej skuteczni w prowadzeniu terapii długoterminowej niż psychologowie, psychiatrzy i pracownicy socjalni.

Ponadto, zdaniem Seligmana kolejnym dowodem na to, że to nie sam upływ czasu jest przyczyną poprawy stanu osób badanych jest fakt, że sami respondenci przyczyn tej poprawy upatrywali w poddaniu się psychoterapii.

Powyższe argumenty Seligmana były wielokrotnie podważane podczas dyskusji nad badaniem CR. Hollon (1996) podkreśla, że do stwierdzenia związku przyczynowego między określonymi oddziaływaniami a zaobserwowanym efektem niezbędna jest obecność grupy kontrolnej, nie zaś ogólne zadowolenie pacjentów z przeprowadzonego leczenia i wiara, że poprawę zawdzięczają właśnie psychoterapii. Hollon nie uważa też, by niższa ocena efektywności psychoterapii przeprowadzanej przez doradców do spraw małżeńskich lub lekarzy ogólnych, czy też związek pomiędzy długością terapii a ogólną z niej satysfakcją stanowiły adekwatne czynniki kontrolne. Również fakt, że dane z badania CR dowodzą równej skuteczności leczenia psychoterapią, jak psychoterapią i lekami nie stanowi wystarczającego dowodu na istnienie specyficznych dla psychoterapii czynników leczących, które sprawiają, że jest ona bardziej skuteczna niż np. rozmowa o swoich problemach z przyjacielem. Jakobson i Christensen (1996) wysuwają bardziej szczegółowe kontrargumenty przeciwko każdemu z twierdzeń Seligmana odnośnie wewnętrznych grup kontrolnych. Ich zdaniem gorsze wyniki terapii przeprowadzanej przez doradców do spraw małżeńskich nie mogą stanowić punktu odniesienia i porównania ze specjalistami w zakresie zdrowia psychicznego, gdyż terapia ta dotyczyła prawdopodobnie innego typu problemów (ograniczona była tylko do problemów małżeńskich), zaś prowadzący ją mogli się w sposób systematyczny (nielosowy) różnić od grupy profesjonalistów. Z kolei lekarze rodzinni prowadzący terapię długoterminową nie mogą być porównywani z profesjonalnymi psychoterapeutami, gdyż ci pierwsi widywali się ze swoimi pacjentami rzadziej. Nieskuteczny jest także, zdaniem autorów, trzeci sposób kontroli. Ma on stanowić początek krzywej obrazującej pozytywną zależność między czasem trwania terapii a jej skutecznością (traktowany jako punkt odniesienia), ponieważ w trakcie terapii długoterminowej spontaniczna remisja jest bardziej prawdopodobna. Wreszcie polemizując z argumentem o równej skuteczności leczenia psychoterapią oraz leczenia psychoterapią i lekami autorzy

ci wskazują, iż nie wiadomo, jakich leków używały osoby biorące udział w badaniu, w związku z jakimi problemami, czy odpowiednio używały tych leków, a także czy osoby badane używające środków farmakologicznych posiadały cechy typowe dla osób, u których leki wykazują lepsze efekty niż placebo.

Wszyscy powoływani wyżej autorzy stoją na stanowisku, że jedynym sposobem, by dowieść związku przyczynowego między oddziaływaniami psychoterapeutycznymi oraz ewentualną poprawą jest zastosowanie grup kontrolnych (porównawczych).

Argumenty krytyczne dotyczące nieobecności grup kontrolnych Seligman (1996a, s. 1075) zaliczył do „nieuzasadnionego krytycyzmu”, niesłusznie usiłującego narzucić badaniom terenowym standardy właściwe dla badań eksperymentalnych. Jego zdaniem, metodą równie poprawną dla dowiedzenia związku przyczynowo-skutkowego między psychoterapią a poprawą stanu osoby leczonej jest model obserwacyjny (*causal modeling*). Seligman twierdzi, że model eksperymentalny pozwala na poszukiwanie związku przyczynowo-skutkowego oraz eliminację przyczyn alternatywnych dzięki zastosowaniu grup kontrolnych oraz przydziału losowego do grup, natomiast model obserwacyjny radzi sobie z tymi problemami poprzez pomiar możliwych przyczyn alternatywnych i sukcesywną ich eliminację. Model eksperymentalny ma tę przewagę nad obserwacyjnym, że eliminuje wpływ zmiennych zakłócających za pomocą jednego zabiegu, a mianowicie randomizacji, jednak z kolei model obserwacyjny pozwala na natychmiastowe odniesienie i generalizację wyniku na sytuacje rzeczywiste, czego nie można zrobić z wynikami uzyskanymi w rezultacie eksperymentu.

Czy rzeczywiście model obserwacyjny jest równie „silny” w wykrywaniu związku przyczynowo-skutkowego? Brzeziński (1996) twierdzi, że niewątpliwie ideałem jest model eksperymentalny, który powinno się stosować w badaniach naukowych tam, gdzie to możliwe (s. 431). Model obserwacyjny uważa on za „słabszy” jeżeli chodzi o jego moc eksplanacyjną, stwarzający możliwość pojawienia się wielu błędów. Podobnego zdania są Cook i Campbell (1979, s. 385), którzy uważają, że: „względny związek z trafnością wewnętrzną badania mają pierwszeństwo, gdy koszt pomyłki związanej z wykryciem związku przyczynowo-skutkowego jest wysoki [...]. Zdecydowanie opowiadamy się za poglądem, że eksperyment respektujący zasadę randomizacji nadal ma większą wartość niż jego alternatywy w postaci *quasi*-eksperymentu czy też metod nieeksperymentalnych, jeżeli chodzi o wykrywanie związku przyczynowo-skutkowego”. Pamiętając zatem o sytuacjach, w których bardzo utrudnione, niemożliwe lub też nieetyczne jest zastosowanie metody „prawdziwego” eksperymentu, należy mieć na względzie owe różnice między „mocą” różnych metodologii w wykrywaniu związku przyczynowo-skutkowego.

Z kwestią obecności grup kontrolnych łączy się także problem randomizacji, czyli losowego doboru osób badanych do grupy eksperymentalnej i kontrolnej. Dzięki temu zabiegowi badacz może kontrolować wpływ zmiennych zakłócających, związanych z osobami badanymi, które zostają w losowy sposób przydzielone do grupy eksperymentalnej i kontrolnej, dzięki czemu zmniejsza się prawdopodobieństwo, że będą one miały systematyczny wpływ na uzyskane wyniki (dzięki np. temu, że uczestnicy grupy eksperymentalnej, składającej się z osób dobranych w sposób nie-losowy, posiadali cechy zwiększające ich podatność na pewnego typu oddziaływania, będące przedmiotem badania) (Brzeziński, 2000b).

Seligman (1995a, s. 973) dowodzi, że co prawda randomizacja zmniejszałaby np. ryzyko tego, że w badaniu mogły wziąć udział osoby, które określa się mianem „fanatyków psychoterapii”, czyli osób, dla których uczestnictwo w terapii jest sposobem na życie i które w związku z tym mają z nią znaczenie bardziej pozytywne doświadczenia. Jednak, zdaniem autora, brak doboru losowego, który miał miejsce w badaniu CR, może być jednym z ważniejszych czynników zwiększających wartość metody stosowanej przez CR i główną wadą metody eksperymentalnej, która nie odzwierciedla prawidłowości zachodzących w rzeczywistości, kiedy to w wielu wypadkach osoba sama poszukuje pomocy terapeuty oraz sama go wybiera, wierząc przy tym, że terapia jej pomoże. Randomizacja w eksperymencie polega także na losowym dobraniu typu psychoterapii do pacjenta i do terapeuty, co także nie oddaje obrazu rzeczywistego – w realnych sytuacjach terapeuta specjalizuje się w pewnej szkole psychoterapeutycznej, która współgra z jego cechami osobowościowymi i pewnymi predyspozycjami, zaś pacjent z określonym typem zaburzeń często sam wybiera, w terapii jakiego typu chce uczestniczyć. Randomizacja może zatem stanowić najsłabszy punkt badań eksperymentalnych, sprawiający, że ich wyniki mogą być dalekie od tego, jak psychoterapia sprawdza się w rzeczywistości, gdzie zamierzone, a nie losowe dopasowanie typu zaburzeń do określonego typu terapii, sprawdzającego się właśnie przy tego typu problemach, wydaje się mieć bardzo istotne znaczenie.

Krytycy Seligmana (Mintz, Drake, Crits-Cristoph, 1996) nie podważają prawdziwości stwierdzenia, iż odpowiednie przydzielenie danej osoby do odpowiedniego typu terapii, odpowiedniego leku i odpowiedniej sekwencji technik terapeutycznych może mieć kluczowe znaczenie dla zaistnienia poprawy stanu pacjenta. Podkreślają oni jednak, że nie jest to argument przeciwko randomizacji – „gdybyśmy wiedzieli, które decyzje są «odpowiednie» (prawidłowe), badania nie byłyby potrzebne” (*op. cit.*, s. 1084).

Należy w pełni zgodzić się z argumentami cytowanych autorów o konieczności zastosowania zewnętrznych grup kontrolnych w badaniu, w którym sprawdza

się istnienie związku przyczynowo-skutkowego pomiędzy przeprowadzeniem terapii a poprawą stanu psychicznego pacjenta. Nawet jeżeli przyjmiemy „łagodniejsze” kryteria poprawności metodologicznej badań terenowych, należy uznać, że jednym z podstawowych mankamentów badania CR jest brak zewnętrznej grupy porównawczej w postaci osób, które deklarowały problemy emocjonalne/psychiczne, ale nie skorzystały z leczenia lub skorzystały z leczenia lub wsparcia nieprofesjonalistów, np. przyjaciół, który to sposób pozwoliłby na określenie stopnia wpływu czynników specyficznych dla psychoterapii i kontrolować wpływ czynników niespecyficznych, czyli zjawisk pojawiających się podczas każdego typu terapii, nie związane z konkretnymi technikami czy założeniami teoretycznymi (należą do nich np. oczekiwania pacjenta związane z wynikiem terapii, nadzieja na osiągnięcie poprawy czy sam kontakt z terapeutą jako człowiekiem, który okazuje pacjentowi uwagę i zrozumienie – Fiske, Hunt, i in., 1970). Tylko obecność grup kontrolnych umożliwia kontrolowanie czynników mających istotny wpływ na trafność wewnętrzną badania, takich jak historia (upływ czasu), związany ze zjawiskiem spontanicznej remisji czy też wpływ czynników niespecyficznych. Czynniki mające stanowić substytut grup kontrolnych, wskazane przez Seligmana, nie mogą zostać uznane za wystarczające.

METODA SAMOOPISU UŻYTA W BADANIU – KWESTIA SUBIEKTYWIZMU I RETROSPEKTYWNOŚCI

W badaniu zastosowano metodę samoopisu (*self-report*). Oznacza to, że wszystkie informacje uzyskane w badaniu CR (w szczególności to, czy rzeczywiście zaszła deklarowana poprawa stanu psychicznego) pochodzą bezpośrednio od osób badanych, a ponadto nie zostały poddane obiektywnej weryfikacji pod kątem ich prawdziwości (Seligman, 1995a, s. 972). Interesuje nas tutaj szczególnie wartość odpowiedzi na trzy pytania zadane w badaniu, a mianowicie dwa pytania opisujące własny stan psychiczny/emocjonalny osoby badanej – na początku terapii oraz w chwili badania i postrzeganą skuteczność psychoterapii.

Fakt braku zewnętrznych źródeł weryfikacji uzyskanych informacji uznaje Seligman za potencjalny, słaby punkt zastosowanej metodologii, lecz broni się przed uznaniem go za element rzeczywistości osłabiający wiarygodność badania wysuwając następujące argumenty:

– Po pierwsze, wszelkie nieścisłości i błędne oceny dotyczące zaszłej zmiany, natury problemu, a nawet rodzaju terapeuty, który przeprowadzał leczenie, mają charakter losowy a nie systematyczny, dlatego też nie stwarzają zagrożenia wprowadzenia – choć tego Seligman nie jest do końca pewien wspominając, że z faktu, iż większość respondentów była prawdopodobnie

w dość dużym stopniu zaangażowana emocjonalnie w psychoterapię, wynika potrzeba przeprowadzenia dalszych badań potwierdzających hipotezę o losowym charakterze wszelkich błędów w informacjach udzielanych przez badanych.

– Drugim, groźniejszym potencjalnym błędem metody samoopisu jest, jak twierdzi Seligman, fakt, że respondenci mogli nie być w stanie prawidłowo opisać swojego stanu psychicznego przed i po leczeniu w zakresie konkretnego problemu, produktywności w pracy i stosunków z innymi ludźmi. Jednak zdaniem autora problem ten dotyka powszechnie nawet doświadczonych diagnostyków, zaś badania wykazują wysoki stopień korelacji między danymi uzyskanymi przy pomocy samoopisu i przy pomocy diagnozy dokonanej przez inną osobę.

Za dość kontrowersyjne należy uznać poglądy analityków *CR* – Kotkina, Daviet i Gurina (1996), którzy broniąc wiarygodności zastosowanej metody samoopisu wysuwają następujące argumenty: „respondenci *CR* w szczególności w znacznym stopniu przypisywali poprawę (którą u siebie zaobserwowali) terapii, z której skorzystali” oraz „choć metoda samoopisu może być związana z pewnymi zniekształceniami (uzyskanych danych), przez lata nauczyliśmy się szanować dane z naszych sondaży i nie lekceważyć danych dostarczanych przez tysiące wykształconych, dokładnych abonentów (raportów *Consumer Reports*), nazywając je «opiniami» lub «dysonansem poznawczym»” (s. 1081).

Wątpliwości Seligmana dzielą Christensen i Jacobson (1996), przypisując im jednak o wiele większą wagę. Ich zdaniem, jednym z zasadniczych problemów zastosowanej metody jest fakt, że w żaden sposób nie można zweryfikować stanów relacjonowanych przez respondentów, nie można ich odnieść do żadnego kryterium obiektywnego. Z kolei inny krytyk badania *CR* – Hollon (1996), oprócz wspomnianych wyżej słabości, podkreśla swe wątpliwości związane z poleganiem na stwierdzeniach osób badanych wskazujących na związek przyczynowo-skutkowy pomiędzy przeprowadzoną terapią a zaobserwowaną poprawą swojego stanu emocjonalnego.

Metoda introspekcji jako źródło pozyskiwania informacji dotyczących stanu psychicznego była różnie oceniana pod względem wiarygodności (Stachowski, 2000). Rosenthal, Rosnow (1991, s. 178) zakładają, że stosowanie metod samoopisu jest uzasadnione wtedy, gdy badacz uważa, że „uczestnicy badania władają językiem oraz posiadają doświadczenie pozwalające im opisywać swoje działania i reakcje”. Zgodnie z tym założeniem, należałoby rozważyć, na ile adekwatnie potrafią ludzie trafnie opisywać swój stan psychiczny, czy potrafią prawidłowo spostrzegać związki przyczynowo-skutkowe, a także jakie czynniki mogą wpłynąć na zniekształcenie informacji udzielanych przez osoby badane.

Rozpatrując kwestię zdolności ludzi do opisywania swojego stanu psychicznego należy stwierdzić, że szczególna cecha respondentów, a mianowicie występowanie u nich zaburzeń natury emocjonalnej, mogła mieć istotny wpływ na zdolność do adekwatnego postrzegania i opisu swego stanu psychicznego. Jeżeli zaś chodzi o postrzegany przez badanych związek psychoterapii z ich obecnym stanem psychicznym można odwołać się do przeglądu badań przedstawionego przez Nisbetta i Wilsona (1995), którzy doszli do wniosku, że chociaż ludzie są (w większym lub mniejszym stopniu) zdolni do opisywania swoich stanów wewnętrznych, wykazują znacznie mniejsze umiejętności w zakresie identyfikowania elementów leżących u podstaw ich reakcji. Uważają oni, że często sprawozdania ludzi dotyczące związku przyczynowo-skutkowego pomiędzy bodźcem a reakcją nie są oparte na introspekcji, a na przyjętych *a priori* ukrytych teoriach przyczynowych, czyli sądach o tym, w jakiej mierze dany bodziec jest prawdopodobną przyczyną danej reakcji. Fakt przypisywania przez respondentów badania *CR* skuteczności oddziaływaniom psychoterapeutycznym może być także w pewnym stopniu efektem zaistnienia zjawiska dysonansu poznawczego opisanego przez Festingera (Aronson, Wilson i Akert, 1997). Wydaje się prawdopodobne, że osoby, które zainwestowały czas, pieniądze i zaangażowanie emocjonalne w swoją terapię, nie będą skłonne do przyznania się, że nie przyniosła ona efektów (Brock, Green i in., 1996).

Można także przypuszczać, że w badaniu mogły pojawić się intencjonalne i nieintencjonalne zniekształcenia informacji udzielanych przez respondentów. Zniekształcenia pierwszego typu pojawiają się często w sytuacjach, gdy poruszane są tzw. kwestie drażliwe, a więc odnoszące się do istotnych, szczególnego rodzaju dziedzin życia osoby badanej. Temat badania *CR* niewątpliwie należał do tego typu tematów, jako że poruszał kwestię zdrowia psychicznego i problemów natury emocjonalnej. Nieobojętność tematu badania może mieć bardzo duży wpływ na udzielane przez respondentów odpowiedzi – klasycznym przykładem tego typu są badania dotyczące alkoholizmu (Mangione, 1999). W tego typu sytuacjach istnieje duże prawdopodobieństwo, że na odpowiedzi wpływ będzie mieć zmienna aprobaty społecznej, która „wypływa ze skłonności osób badanych do nieprzyznawania się do cech społecznie niepożądanych i z udawania, że się ma cechy korzystne oraz z tendencji do mówienia rzeczy stawiających je w korzystnym świetle” (Nederhof, 1991, s. 212). Można zatem przypuszczać, że zmienna aprobaty społecznej mogła wywrzeć wpływ na odpowiedzi osób badanych pytanych o dane dotyczące ich stanu psychicznego – przyznawanie się do problemów i zaburzeń psychicznych nie jest zjawiskiem pożądanym społecznie, a respondenci mogli kierować się chęcią podtrzymywania dobrego wrażenia o sobie (Nederhof, *op. cit.*).

Odpowiedzi osób badanych mogły także ulec zniekształceniu nieintencjonalnemu – należy tu wspomnieć o wpływie zapominania czy też nieadekwatnej oceny sytuacji przeszłej, tu: swego samopoczucia przed podjęciem terapii, z perspektywy terażniejszości.

Innym potencjalnym źródłem zniekształcenia wypowiedzi respondentów mógł być wpływ wskazówek sugerujących hipotezę badawczą (Orne, 1991). Respondentom nietrudno było zgadnąć, że badanie *CR* jest badaniem mającym na celu weryfikację hipotezy o skuteczności psychoterapii. Nie można zatem wykluczyć, że respondenci ochotniczo biorący udział w badaniu wykazali zachowania typowe dla „dobrych osób badanych”, które będąc same przekonane o słuszności jakiejś tezy pragną dopomóc w dowiedzeniu jej przez badacza.

Należy także pamiętać o zagrożeniach dla wiarygodności danych, których źródłem jest fakt zastosowania w badaniu metody retrospektywnej, a więc opierającej się na wspomnieniach osób badanych odnośnie ich stanu emocjonalnego przed terapią. Mimo oczywistych wad tej metody, ma ona zdaniem Seligmana jedną zaletę – zwiększa trzeźwość oceny, gdyż po pewnym czasie od zakończenia terapii, gdy opadły już euforia i idealizacja, ma się do niej większy dystans (Seligman, 1995a). Choć zdaniem autora zastosowanie metody prospektywnej byłoby dużo lepsze, to jednak metoda retrospektywna nie jest istotną wadą metodologiczną badania, gdyż jest prawdopodobnie źródłem błędów o charakterze losowym, a nie systematycznym. Jednak to ostatnie twierdzenie wydaje się jedynie hipotezą – sam Seligman postuluje przeprowadzenie badań mających na celu potwierdzenie jej. Niemniej jednak autor dowodzi, że zastosowana metoda jest poprawna metodologicznie i nie dostarcza błędnych danych – powołuje się na liczne badania przeprowadzone wcześniej przez *CR*, w których różne produkty badane były dwutorowo – przy pomocy eksperymentów laboratoryjnych oraz przy pomocy ankiet odwołujących się do opinii respondentów na temat tych produktów. Jako że dane uzyskane przy pomocy obu metod były spójne, metoda ankietowa uznana została za prawidłową. I tu ponownie nie można zgodzić się z argumentami Seligmana. Fakt, że przedmiotem badania był stan psychiczny respondentów oraz jego związek z odbytą psychoterapią sprawia, że nie można go porównywać z badaniami przedmiotów codziennego użytku, jako że w tym drugim przypadku mało prawdopodobne jest wystąpienie zniekształceń, które mogą pojawić się w badaniu dotyczącym temat drażliwy, dotyczący respondenta osobiście. Dlatego też nie można przyjąć jako oczywistego założenia o losowym rozkładzie ewentualnych zniekształceń ocen osób badanych, mając na względzie wszelkie możliwe źródła zniekształceń uzyskanych danych opisane w poprzednim podrozdziale.

ZASTOSOWANE NARZĘDZIE BADAWCZE – KWESTIONARIUSZ

Narzędziem zastosowanym w badaniu był kwestionariusz do samodzielnego wypełniania, w którym respondenci wybierali odpowiedź z podanej kafeterii. Nie podlega dyskusji fakt, że sposób sformułowania pytań oraz kolejność ich zadawania ma bardzo istotny wpływ na rodzaj uzyskiwanych odpowiedzi. Niezrozumienie lub błędne zrozumienie pytania powoduje znaczne zniekształcenia uzyskiwanych danych (Johnson, 1991, Mangione, 1999, Tyszką, 1995).

Warto w tym miejscu odwołać się do danych przytaczanych przez Tyszkę (1995, s. 16-17), który cytuje wyniki badania satysfakcji konsumentów z różnego rodzaju produktów i usług. W badaniu tym okazało się, że typowym oszacowaniem jest „bardzo usatysfakcjonowany”, zaś rozkład zadowolenia jest silnie skośny w kierunku wysokiego poziomu zadowolenia. Autorzy badania (Peterson i Willson, 1992, za: Tyszką, *op. cit.*) zauważyli jednak, że wystarczy zmienić sformułowanie pytania i zamiast zadać je w formie „jak bardzo jesteś usatysfakcjonowany (czymś)?” na „jak bardzo jesteś nieusatysfakcjonowany (czymś)?”, by pokazać „zmniejszyć” odsetek zadowolonych konsumentów. Przykład ten można odnieść do zadanego w badaniu *CR* pytania o satysfakcję z psychoterapii – nie można wykluczyć, że zmiana sformułowania tego pytania mogłaby mieć istotny wpływ na udzielone odpowiedzi.

Warto też zauważyć, że wśród pytań zadanych w badaniu znalazły się między innymi pytania o typ podjętej terapii (terapia psychodynamiczna, behawioralna, poznawcza, feministyczna) oraz o kompetencje terapeuty. Należy uznać, że zadawanie tego typu pytań jest poważnym błędem metodologicznym, jako że opiera się na fałszywym moim zdaniem założeniu, że osoba biorąca udział w terapii dysponuje fachową wiedzą na temat rodzaju terapii oraz kompetencji terapeuty. Innymi słowy, pytania te odwołują się do wiedzy, której respondent prawdopodobnie nie posiada lub którą posiada w niewystarczającym zakresie. Ponadto należy mieć na względzie, że bardzo popularnym typem terapii jest terapia eklektyczna, łącząca elementy różnych szkół psychoterapeutycznych, której istnienie zupełnie zignorowali autorzy badania.

Bardziej szczegółowa analiza konstrukcji kwestionariusza i sposobu zadawania pytań, która mogłaby zidentyfikować inne potencjalne źródła zniekształceń uzyskanych danych nie jest niestety możliwa, jako że ani *Consumer Reports* ani *Consumer Union* nie chciały udostępnić autorce artykułu kwestionariusza wykorzystanego w badaniu.

Brak jest również danych odnośnie rzetelności narzędzia, która to kwestia została podniesiona przez Jackobsona (1995). Przypomnijmy, że rzetelność narzędzia pomiarowego definiowana jest jako jego zdolność

do dokładnego pomiaru danej zmiennej, niezależnie od tego, co w rzeczywistości mierzy dane narzędzie (Brzeziński, 1996, Magnusson, 1981). W odpowiedzi na pytanie Jacobsona, Seligman (1995b) stwierdził, że „rzetelność (narzędzia) jest całkiem niezła, z tego, co pamiętam (nie mam przy sobie dokładnych danych)”. Brak jest poza tym jakichkolwiek danych dotyczących rzetelności kwestionariusza użytego w badaniu.

ZASTOSOWANE MIARY WYNIKU – SPOSÓB DEFINIOWANIA POPRAWY STANU PACJENTA

Częstym zarzutem wobec badania *CR* jest argument, że nie można go uważać za badanie efektywności psychoterapii, a jedynie za sondę mierzącą poziom satysfakcji osób, które korzystały z psychoterapii. Seligman wielokrotnie (1995b, 1996c) zaprzeczał, jakoby badanie *CR* było jedynie badaniem satysfakcji osób, które poddały się leczeniu psychoterapią. Dowodzi on, że zadowolenie z terapii (pomiar retrospektywny – ocena satysfakcji własnej osoby badanej po przeprowadzeniu terapii) było tylko jednym z trzech wskaźników mierzonych w badaniu. Ponadto, pomiarowi poddano dwa inne wskaźniki, a mianowicie:

- poprawa specyficzna – polepszenie stanu osoby w zakresie konkretnego problemu zgłoszonego do terapii (pomiar retrospektywny – opinia osoby badanej po przeprowadzeniu terapii na temat zaszłej zmiany)
- poprawa ogólna – ogólne polepszenie funkcjonowania osoby badanej (miarą polepszenia była różnica między opisywanym retrospektywnie stanem osoby badanej przed i po terapii).

W badaniu pytano respondentów również o ocenę poprawy ogólnego funkcjonowania pod względem efektywności w pracy, jakości związków interpersonalnych, samopoczucia, możliwości wglądu i rozwoju.

Pewien problem zdaniem Seligmana (1995a) może stanowić fakt, iż zastosowane wskaźniki poprawy miały charakter molarne, czyli dość ogólny – badanych pytano np. o ocenę, na ile terapia pomogła im w rozwiązaniu tych problemów, które skłoniły ich do podjęcia leczenia. Autor twierdzi, że bardziej precyzyjne byłoby użycie wskaźników molekularnych, czyli mierzących szczegółowo konkretne zachowania będące obserwowalnymi przejawami np. poprawy, o której mowa w ogólnym pytaniu. Zarzut ten może być, zdaniem Seligmana, sformułowany jako „brak czułości” użytych metod pomiarowych, które to podejrzenie może być potęgowane przez brak zarejestrowanych różnic między różnymi rodzajami terapii. Kolejnym zarzutem wobec zastosowanych miar wyników badania może być ten dotyczący błędnych opisów odpowiedzi na skalach szacunkowych, których wartość uzależniona jest w pełni od ich wartości fasadowej, brak im zaś odniesienia do norm psychometrycznych. Niemniej jednak podejście

„molarne” może znaleźć pewne uzasadnienie, którym jest pojęcie istotności klinicznej zmiany. W literaturze częstokroć (Hansen i Lambert, 1996, Jacobson i Truax, 1991) podkreślano, że nie wystarczy oprzeć się na pojęciu istotności statystycznej zaobserwowanej zmiany – z punktu widzenia celów psychoterapii znaczenie ma istotność kliniczna, która odnosi się do pojęcia praktycznej wagi zmiany osiągniętej dzięki psychoterapii. Kazdin (1977) stwierdza, że „zmiany ważne z punktu widzenia klinicznego powinny być dramatyczne i w oczywisty sposób widoczne w posiadanych danych, tak że nie istnieje potrzeba odwoływania się do testów statystycznych” (s. 428). Sugeruje on, by w ocenie istotności zmiany odnieść się do metod walidacji społecznej, które dokonują osądu wyniku psychoterapii w oparciu o normy społeczne. Metoda pomiaru istotności zmiany zastosowana w badaniu *CR* wydaje się zatem nawiązywać do tez Kazdina dotyczących „widoczności zmiany gołym okiem”. Seligman (1995a, s. 970) twierdzi, że badanie *CR* poradziło sobie z tym problemem w dość prosty sposób – respondenci odpowiadali wprost na pytanie o to, na ile terapia była skuteczna w leczeniu ich problemów (krańce skali opisane były stwierdzeniami „znacznie poprawiła” i „znacznie pogorszyła”). Twierdzenia te mają wykazywać rzeczywistą istotność zmiany, którą zaobserwowały osoby ją opisujące.

Należy jednak zwrócić uwagę na fakt, że ocena poprawy stanu osoby poddanej terapii dokonywana była z jednej tylko perspektywy, a mianowicie samego pacjenta. W tym miejscu warto powołać się na koncepcję zdrowia psychicznego sformułowaną przez Hadleya i Struppa (1977), w której pojęcie to ujmowane jest z trzech perspektyw – perspektywy społeczeństwa, pacjenta oraz specjalisty do spraw zdrowia psychicznego. Każdy z tych trzech podmiotów definiuje zdrowie psychiczne w specyficzny, odmienny sposób, uwzględniający właściwe swej perspektywie cele, jakim ma służyć ta definicja. Autor postuluje zatem, by adekwatna, wszechstronna ocena zdrowia psychicznego danej osoby oraz zmian, które zaszły w wyniku psychoterapii uwzględniała wszystkie trzy wspomniane perspektywy i wskazywała na ewentualne rozbieżności. Ograniczenie się do użycia problemu z jednej tylko perspektywy powinno być, zdaniem Struppa, jasno podkreślone, wraz z zastrzeżeniem ograniczeń tego podejścia, jako że ocena wyniku terapii zależy w dużej mierze od perspektywy osoby dokonującej oceny (Strupp, 1986). Na istotność przyjęcia perspektywy społeczeństwa dla oceny istotności zmiany zaszłej w wyniku psychoterapii wskazywali także Jacobson i Truax (1991), których koncepcja za istotną uznaje zmianę polegającą na przesunięciu osoby, która przed terapią należy do grupy osób dysfunkcyjnych („pacjentów”) do grupy osób „normalnych”. Kryterium oceny społecznej, do którego odwołuje się ta koncepcja, mimo swej stronniczości wydaje się bardziej adekwatne

niż kryterium zadowolenia pacjenta z psychoterapii i jego przekonania o skuteczności leczenia.

Seligman (1996a) podzielił opinię Struppa (1996) na temat jednostronności zastosowanej perspektywy postrzegania poprawy będącej wynikiem psychoterapii. Seligman przyznaje, że lepszym rozwiązaniem byłoby zastosowanie oprócz metody samoopisu oceny stanu pacjenta przez terapeutę oraz użycie baterii narzędzi diagnostycznych. Niemniej jednak autor w konkluzji swych wszystkich wywodów nie przyznaje, że ocena zmiany stanu pacjenta przypisywana przez niego psychoterapii jest zbyt słaba, by podważyć ogólny wniosek z badania o skuteczności psychoterapii w ogóle.

Podstawowymi wadami metodologicznymi przyjętego sposobu ujęcia zmiany zaszłej w wyniku psychoterapii są zatem:

- nieprecyzyjne narzędzie pomiarowe, opierające się jedynie na ogólnych stwierdzeniach dotyczących zauważonej poprawy;
- przyjęcie jednostronnej perspektywy – ocena zmiany jedynie przez osobę leczoną, bez uwzględnienia opinii społeczeństwa czy ekspertów;
- oparcie najważniejszego wniosku z badania: „psychoterapia jest skuteczna” na ogólnym wskaźniku efektywności terapii będącym pochodną m.in. wskaźnika satysfakcji z psychoterapii, a zatem miary wysoce subiektywnej i nie muszącej mieć nic wspólnego z rzeczywistymi rezultatami terapii. Dane przytaczane przez Lish (1995, badanie McLellana nad efektywnością leczenia alkoholizmu i narkomanii) wskazują na praktycznie zerową korelację między satysfakcją pacjenta z przeprowadzonego leczenia a jego wynikami.

REGRESJA DO ŚREDNIEJ A WYNIKI CR

Wyniki badania wskazują, że psychoterapia była niezwykle skuteczna wobec całego spektrum różnych typów terapii i różnych typów zaburzeń (Seligman, 1995a). Z 426 osób, które czuły się „bardzo źle” przed rozpoczęciem terapii, 87% czuło się „bardzo dobrze”, „dobrze” lub przynajmniej „tak sobie” w trakcie przeprowadzania badania. Z 786 osób, które twierdzą, że czuły się „raczej źle” na początku leczenia, 92% czuło się „bardzo dobrze”, „dobrze” lub przynajmniej „tak sobie” w czasie, kiedy przeprowadzano badanie (*op. cit.*, s. 968).

Fakt ten podnieśli m. in. Brock, Green i in. (1996), którzy zasugerowali, że wyniki te mogą być związane nie ze skutecznością zastosowanego leczenia, lecz ze zjawiskiem regresji do średniej (regresji statystycznej). Zjawisko to polega na tendencji do zbliżania się skrajnych wyników uzyskanych w pre-teście do średniej w post-teście (Brzeziński, 2000b), która to tendencja nie jest efektem czynnika eksperymentalnego, czyli w tym przypadku poddania się leczeniu. Regresja staty-

styczna jest istotnym czynnikiem obniżającym trafność wewnętrzną badania, zaś badacz prowadzący badanie metodą eksperymentalną powinien unikać brania pod uwagę osób o skrajnych wynikach (*ibidem*).

WNIOSKI

Przedstawiony wyżej przegląd zagadnień związanych z metodologią badania efektywności psychoterapii przeprowadzonego przez *Consumer Reports* dowodzi istnienia w badaniu poważnych uchybień i braków metodologicznych. Brock, Green i in. (1996) określają je mianem „klasycznych” wręcz uchybień metodologicznych, które – ku ich niepomierzemu zdziwieniu – Seligman nie tylko zbagatelizował, ale przedstawił wręcz jako zalety. Podsumowując, należy stwierdzić, że podstawowymi brakami podważającymi wiarygodność uzyskanych wyników są stroniczy dobór próby, brak grup kontrolnych oraz dość nieprecyzyjne i jednostronne narzędzie pomiaru. Każde z tych uchybień z osobna wystarcza, by uznać wyniki badania CR za bezwartościowe, cechujące się zarówno niską trafnością wewnętrzną, jak i zewnętrzną, bowiem zastosowana metodologia nie pozwoliła na kontrolę czynników ograniczających oba rodzaje trafności, takich jak: historia, dojrzwienie, selekcja, regresja statystyczna, utrata osób badanych w przypadku trafności wewnętrznej oraz niereprezentatywności próby oraz interakcji selekcji z czynnikiem aktywnym (terapią) w przypadku trafności zewnętrznej. Nie można się zatem w tym przypadku oprzeć na tezie wskazującej na konieczność rezygnacji z rygorystycznych warunków przeprowadzenia badania na wzór badania eksperymentalnego kosztem trafności jego wewnętrznej na rzecz zwiększenia trafności zewnętrznej (*efficacy vs. effectiveness study trade-off*, Seligman, 1996a, s. 1076), jako że w badaniu CR nie może być mowy o jego trafności zewnętrznej, którą niweczy całkowicie stroniczy dobór próby.

Kolejne argumenty Seligmana broniącego zastosowanej metodologii i umniejszającego wagę kolejnych jej słabości wydają się być nieprzekonujące. Należy przy tym podkreślić, że mimo słusznie podniesionego przez Seligmana argumentu dotyczącego istotnych różnic między „wysterylizowanym” eksperymentem laboratoryjnym a badaniem naturalistycznym, trzeba mieć na względzie podstawowe warunki, które musi spełniać nawet badanie prowadzone w warunkach naturalnych, jeżeli ma być określone mianem naukowego. Mintz, Drake i Crits-Christoph (1996, s. 1084) są zdania, że „prawdziwe badanie efektywności psychoterapii metodami obserwacyjnymi jest tak samo rygorystyczne metodologicznie, jak eksperymenty laboratoryjne, jest narażone na takie same zagrożenia trafności wewnętrznej i zwykle ociera się o wiele tych samych fundamentalnych problemów metodologicznych”.

Autorzy ci wskazują na następujące warunki, które muszą być przestrzegane podczas przeprowadzania badań typu *effectiveness*: 1) dokładna charakterystyka okoliczności prowadzenia terapii oraz terapeuty, 2) reprezentatywność próby poddanej badaniu, 3) określenie typu i „ilości” uzyskanej przez pacjenta terapii, 4) ocena przyczyn braku uczestnictwa w terapii, i najważniejszy zdaniem autorów warunek 5) obecność zewnętrznych grup kontrolnych zachowujących przynajmniej warunki stawiane badaniom *quasi*-eksperymentalnym, jednakże z preferencją zastosowania randomizacji. Autorzy twierdzą, że dochowanie tych wszystkich warunków nie jest niemożliwe, wskazując na szereg badań, które ich przestrzegały (*op. cit.*, s. 1084).

Fakt, że badanie *CR* nie było badaniem typu eksperymentalnego, ani też *quasi*-eksperymentalnego, jako że badacz nie miał w nim żadnego wpływu na manipulację zmiennymi stanowi o jego nieprzydatności jako sposobu wykrycia związku przyczynowo-skutkowego (Brzeziński, 2000a). Należy zatem uznać, że główny wniosek badania *CR*, który zawiera się w stwierdzeniu „psychoterapia jest skuteczna” jest pozbawiony podstaw naukowych, jako że zastosowana metodologia nie pozwala na sformułowanie takiego wniosku. Warto przy tym pokreślić, że wielu autorów (Jacobson i Christensen, 1996) kwestionuje użyteczność stawiania pytania o skuteczność psychoterapii „w ogóle” wobec problemów psychicznych traktowanych ogólnie, postulując pomiar skuteczności konkretnych typów interwencji w leczeniu konkretnych zaburzeń.

Badanie *CR* nie spełnia także warunków stawianego badaniom typu *effectiveness*, które mają za zadanie mierzyć użyteczność terapii w warunkach rzeczywistych i określić rozmiar jej efektów (Nathan, Stuart i Dolan, 2000). Także w tym zakresie metodologia badania *CR* nie pozwoliła na zebranie wiarygodnych danych, w związku z brakiem reprezentatywności próby oraz nieprecyzyjnym pomiarem zarówno stanu pacjentów przed i po terapii, jak również brakiem szczegółowych danych na temat zastosowanej terapii.

Zalety badania *CR*, takie jak jego często podkreślany naturalistyczny charakter *CR* polegający na podobieństwie sytuacji i osób w nim badanych do sytuacji i osób rzeczywistych był często podkreślaną zaletą badania *CR* (Seligman, 1995a, 1996a; Kotkin, Daviet, Gurin, 1996), jak również niskie koszty, które to czynniki miały stanowić o jego przewadze nad eksperymentami laboratoryjnymi nie równoważąc niskiej wartości metodologicznej badania *CR*, która podważa wartość uzyskanych w nim wyników.

Mimo tego krytycznego wniosku, należy podkreślić istotną zaletę badania *CR*, a mianowicie wagę, jaką przyłożono w nim do uwzględnienia perspektywy samego pacjenta, która jest często zaniedbywana w badaniach nad psychoterapią. Mcleod (1998, za: Gordon, 2000) twierdzi, że bardzo niewielu badaczy próbowało

przeprowadzić badania, w których sami pacjenci byli pytani o swoje odczucia wobec terapii, w której uczestniczyli. Autor ten wskazuje na silną presję instytucji, zachęcających naukowców do prowadzenia badań nad psychoterapią w sposób naśladowujący nauki ścisłe, a więc stosowania badań w kontrolowanych warunkach eksperymentalnych, stosujących obiektywne miary wyniku, wywiedzione z modeli teoretycznych. Z tej perspektywy subiektywne odczucia jednostki, jej opinie nie leżą w kręgu zainteresowania badaczy. Należy jednak zgodzić się z poglądem (Gordon, *op. cit.*, Macran i in., 1999), że uwzględnienie punktu widzenia pacjenta ma istotne znaczenie dla pełniejszego zrozumienia procesu psychoterapii i jego wyniku. Niektórzy badacze (badania Cartwrighta, Kirtnera i Fiske (1963) Struppa, Foxa i Lesslera (1969) cytowane przez Lamberta i Bergina, 1991, s. 290) podjęli już pewne próby w tym kierunku stosując metody takie, jak Kwestionariusz Post-terapeutyczny Pacjenta, zawierające pytania o zadowolenie pacjenta z leczenia, zauważone zmiany, itp. Zaletami tego typu metod wskazywanymi przez autorów są prostota i szybkość, jednak stanowią one jedynie „grube” narzędzie oceny skuteczności terapii.

Konkludując, należy docenić wkład badania *CR* jako badania zwracającego się bezpośrednio i na dużą skalę do odbiorców psychoterapii i podkreślającego wagę ich opinii na temat leczenia. Tym samym badanie to zrealizowało częściowo postulat Struppa (1986, 1996) czy też Lamberta i Bergina (1991), by w badaniach nad psychoterapią uwzględniać różne perspektywy oceny, co zapobiega systematycznym, tendencyjnym zniekształceniom i w związku z tym także nieuzasadnionym wnioskom. Należy jednak pamiętać o licznych uchybieniach metodologicznych badania *CR*, które sprawiają, że jego wyników nie można uznać za uzasadnione naukowo.

Czy badanie efektywności psychoterapii mogło być przeprowadzone w sposób, który nie zniwelowałby wysiłków badaczy biorących udział w tym wielkim przedsięwzięciu? Sam Seligman (1995a) przyznaje, że idealne badanie typu *effectiveness* powinno połączyć naturalizm badań kwestionariuszowych z zaletami badań eksperymentalnych, a więc: uwzględniać obecność grup kontrolnych, mieć charakter prospektywny, w którym duża próba osób poszukujących terapii (cierpiących na więcej niż jedno zaburzenie i mających możliwość wyboru terapii) zostałaby poddana szczegółowym badaniom przed i po terapii (której czas trwania byłby uzależniony od poprawy pacjenta i która miałaby charakter samokorekcyjny). Jako narzędzie używane w badaniu Seligman proponuje kwestionariusz zawierający szczegółowe opisy zachowań osoby badanej oraz pytania dotyczące stopnia ogólnej poprawy funkcjonowania. Po trzecie, za korzystne uznaje autor włączenie ślepej diagnostycznie oceny funkcjonowania badanych, która zmniejszyłaby zniekształcenia spowodowane

subiektywizmem respondentów.

Do propozycji Seligmana należało by z pewnością dołączyć postulat losowego doboru próby osób biorących udział w badaniu, pomiaru innych czynników mogących mieć wpływ na poprawę stanu pacjenta oraz zastosowania pomiaru z innej perspektywy niż tylko perspektywy pacjenta – jako że dotychczasowe badania dowodzą braku zgodności pomiędzy poszczególnymi metodami pomiaru, warto posługiwać się wskaźnikami reprezentującymi różne punkty widzenia. Czy jednak dokładając kolejne elementy powszechnie uznawane za konieczne dla poprawności metodologicznej badania obserwacyjnego typu *effectiveness*, mamy możliwość wykrycia związku przyczynowo-skutkowego między interwencją terapeutyczną a poprawą w poziomie funkcjonowania pacjenta? Na pytanie to należy odpowiedzieć negatywnie. Eksperyment pozostaje jedynym sposobem wykrycia tego typu związku i – co należy podkreślić – nie musi być on obarczony wszystkimi słabościami wskazanymi przez Seligmana, a przede wszystkim odległością warunków i sposobu przeprowadzenia terapii od warunków rzeczywistych.

Jednym z rozwiązań wydaje się być metoda eksperymentu, albo przynajmniej *quasi*-eksperymentu przeprowadzanego w otoczeniu naturalnym (*clinical trial*) (Jacobson i Christensen, 1996, West, 1997, Hollon, 1996). Taka wersja eksperymentu pozwala na pewną elastyczność traktowania poszczególnych elementów poddanych badaniu, przy jednoczesnym poddaniu ich ścisłej kontroli pod kątem zagrożeń trafności wewnętrznej i zewnętrznej. Należy zatem zgodzić się z Clarke (1998), który zarzuca rozgraniczeniu *efficacy* – *effectiveness* pewną sztuczność i postuluje połączenie elementów badań typu laboratoryjnego i badań typu naturalistycznego, a więc eksperyment naturalny przeprowadzany w warunkach zbliżonych do naturalnych, jednocześnie przestrzegający warunków rygorystycznej kontroli właściwej badaniom eksperymentalnym.

LITERATURA

- Aronson E., Wilson T.D., Akert R.M. (1997). *Psychologia społeczna. Serce i umysł*. Poznań: Wyd. Zysk i S-ka.
- Barlow, D.H. (1996). Health care policy, psychotherapy research, and the future of psychotherapy. *American Psychologist*, 51, 1050-1058.
- Bergin, A. E., Lambert, M. J. (1991). Ocena wyników terapii. W: S. L. Garfield, E. A. Bergin, (red.) *Psychoterapia i zmiana zachowania* (s. 231-308). Warszawa: Instytut Psychiatrii i Neurologii.
- Brock, T., Green, M., Reich, D., Evans, L. (1996). The Consumer Report study of psychotherapy: Invalid is invalid. *American Psychologist*, 51, 1083.
- Brock, T., Green, M., Reich, D. (1998). New evidence of flaws in the Consumer Reports study of psychotherapy. *American Psychologist*, 52, 62-63.
- Brzeziński, J. (1996). *Metodologia badań psychologicznych*. Warszawa: Wyd. Nauk. PWN.
- Brzeziński, J. (2000a). Podstawowe modele badawcze: eksperymentalny i korelacyjny. W: J. Strelau (red.), *Psychologia. Podręcznik akademicki* (t. I, s. 335-387). Gdańsk: Gdańskie Wyd. Psychologiczne.
- Brzeziński, J. (2000b). *Badania eksperymentalne w psychologii i pedagogice*. Warszawa: Wyd. Nauk. Scholar.
- Clarke, G. N. (1998). Improving the transition from basic efficacy research to effectiveness studies: methodological issues and procedures. W: A. Kazdin (red.), *Methodological Issues & Strategies in Clinical Research* (s. 541-560). Washington, D.C.: American Psychological Association.
- Cook, T., Campbell, D. (1979). *Quasi-Experimentation. Design & analysis for field settings*. Boston: Houghton Mifflin Company.
- Fiske, D., Hunt, H., Lubrosky, L., Orne, M., Parloff, M., Reiser, M., Tuma, A. (1970). The planning of research on effectiveness of psychotherapy. *American Psychologist*, 25, 727-737.
- Gordon, N. S. (2000). Researching psychotherapy, the importance of the client's view: a methodological challenge. *The Qualitative Report*, 3 i 4.
- Hansen, N., Lambert, M. (1996). Clinical significance. An overview of methods. *Journal of Mental Health*, 5, 17-24.
- Hollon, S. (1996). The efficacy and effectiveness of psychotherapy. Relative to medications. *American Psychologist*, 51, 1025-1030.
- Howard, K.L., Moras, K., Brill, P., Martinovich, Z., Lutz, W. (1996). Evaluation of Psychotherapy. Efficacy, effectiveness and patient progress. *American Psychologist*, 51, 1059-1064.
- Jacobson, N. (1995). Głos w dyskusji w: *The Consumer Report*. Wypowiedzi z internetowej listy dyskusyjnej SSCPNET 6/95.
- Jacobson, N., Christensen, A. (1996) Studying the effectiveness of psychotherapy. How well can clinical trials do the job? *American Psychologist*, 51, 1031-1039.
- Jacobson, N., Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19.
- Johnson, R. (1991). Pułapki w badaniu: wywiad jako model przykładowy. W: J. Brzeziński, J. Siuta (red.), *Społeczny kontekst badań psychologicznych i pedagogicznych* (s. 433-450). Poznań: Wyd. Nauk. UAM.
- Kotkin, M., Daviet, C., Gurin, J. (1996) The Consumer Reports Mental Health Survey. *American Psychologist*, 51, 1080-1082.
- Kotkin, M., Daviet, C. (1999) Mistaken inferences. *American Psychologist*, 54, 1130.
- Lipsey, M., Wilson, D. (1993). The efficacy of psychological, educational, and behavioural treatment. *American Psychologist*, 48, 1181-1209.
- Lish, J. (1995). Głos w dyskusji w: *The Consumer Report*. Wypowiedzi z internetowej listy dyskusyjnej SSCPNET 6/95.
- Macran, S., Ross, H., Hardy, G. E., Shapiro, D.A. (1999). The importance of considering clients' perspectives in psychotherapy research. *Journal of Mental Health*, 8, 325-337.
- Magnusson, D. (1981). *Wprowadzenie do teorii testów*. Warszawa: Wyd. Nauk. PWN.
- Mangione, T. (1999). *Ankietowanie pocztowe w badaniach marketingowych i socjologicznych*. Warszawa: Wyd. Nauk. PWN.

- Mintz, J., Drake R., Crits-Christoph P. (1996) Efficacy and effectiveness of psychotherapy: two paradigms, one science. *American Psychologist*, 51, 1084-1085.
- Nathan, P. E., Stuart, S. P., Dolan S. L. (2000). Research on psychotherapy efficacy and effectiveness: Between Scylla and Charybdis? *Psychological Bulletin*, 6, 964-981.
- Nederhof, A. (1991). Metody korygowania zniekształcającego wpływu zmiennej aprobaty społecznej: przegląd badań. W: J. Brzeziński, J. Siuta (red.). *Spółeczny kontekst badań psychologicznych i pedagogicznych* (s. 211-234). Poznań: Wyd. Nauk. UAM.
- Nisbett, R.E., Wilson, T. D. (1995). Powiedzieć więcej, niż się wie. Sprawozdania słowne o procesach psychicznych. W: T. Tyszka (red.), *Czy powrót do introspekcji?* (s. 59-115). Warszawa: Wyd. Nauk. PWN.
- Orne, M. (1991). Eksperyment psychologiczny z punktu widzenia psychologii społecznej ze szczególnym uwzględnieniem wpływu zmiennych sugerujących hipotezę badawczą i ich implikacji. W: J. Brzeziński, J. Siuta (red.). *Spółeczny kontekst badań psychologicznych i pedagogicznych* (s. 17-32). Poznań: Wyd. Nauk. UAM.
- Rosenthal, R., Rosnow, R. (1991). *Essentials of Behavioural Research. Methods and Data Analysis*. New York: McGraw Hill, Inc.
- Seligman, M. (1995a). The effectiveness of psychotherapy: The Consumer Reports study. *American Psychologist*, 50, 965-974.
- Seligman, M. (1995b). Głos w dyskusji w: *The Consumer Report*. Wypowiedzi z internetowej listy dyskusyjnej SSCPNET 6/95.
- Seligman, M. (1996a). Science as an ally of practice. *American Psychologist*, 51, 1072-1079.
- Seligman, M. (1996b) Skuteczność psychoterapii. Badania ankietowe Consumer Reports. *Czasopismo Psychologiczne*, 2, 101-112.
- Seligman, M. (1996c). A creditable beginning. *American Psychologist*, 51, 1086-1088.
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32, 752-760.
- Stachowski, R. (2000). *Historia współczesnej myśli psychologicznej. Od Wundta do czasów najnowszych*. Warszawa: Wyd. Nauk. Scholar.
- Strupp, H. (1986). Psychotherapy. Research, practice and public policy (how to avoid dead ends). *American Psychologist*, 41, 120-130.
- Strupp, H. (1996). The tripartite model and the Consumer Reports Study. *American Psychologist*, 51, 1017-1024.
- Strupp, H., Hadley, S. (1977). A tripartite mode of mental health and therapeutic outcomes: With special reference to the negative effects of psychotherapy. *American Psychologist*, 32, 187-196.
- Tyszka, T. (1995). Wprowadzenie. Dawne i nowe spory wokół introspekcji. Czy powrót do introspekcji? W: T. Tyszka (red.), *Czy powrót do introspekcji?* (s. 9-30). Warszawa: Wyd. Nauk. PWN.
- Wampold, B. E., (1997). Methodological problems in identifying efficacious psychotherapies. *Psychotherapy Research*, 7, 21-43.
- West, S. G. (1997). Poza eksperyment laboratoryjny – plany eksperymentalne oraz *quasi*-eksperymentalne w otoczeniu naturalnym. W: I. Heszen-Niejodek, H. Sęk (red.), *Psychologia zdrowia* (s. 62-174). W-wa: Wyd. Nauk. PWN.