

Skala dwuważonych ocen

Skala dwuważonych ocen (ang. *Two-weighted Ratings Scale* – TwRS) – wzorzec strategii pomiarowej i rodzaj narzędzia służącego do badań społecznych, dedykowanego zwłaszcza na potrzeby *ewaluacji splotowej*. Głównym kryterium budowania indeksów takiej skali jest *konkluzyjność* (wnioskotwórczość) wskaźników odzwierciedlających rzeczywistość. Wymaga to dwuwymiarowego ujęcia różnych aspektów badanych zjawisk w wymiarach osobowych, czasowych i przestrzennych, co osiąga się poprzez celowe zróżnicowanie treści bodźców-stwierdzeń w kwestionariuszu. Nazwę *Skali dwuważonych ocen* uzasadnia także integralny sposób analiz statystycznych, które polegają na porównywaniu wartości i ważkości wypowiedzi respondentów^[1].

Spis treści

[Założenia jakościowe](#)

[Założenia kwantytatywne](#)

[Założenia standaryzacyjne](#)

[Zobacz też](#)

[Przypisy](#)

[Bibliografia](#)

Założenia jakościowe

- [Operacjonalizacja](#)^[2] jako nadawanie sensu empirycznego wskaźnikom jakościowym,
- [Strategia psychometryczna](#) w budowaniu pośrednich próbników i wskaźników cech^[1].

Bodźce-stwierdzenia

Przejście z pojęć na poziom empiryczny wymaga zbudowania *skali pomiarowej*. W tym modelu zamiast pytań stosujemy stwierdzenia wyrażające mikrotezy badawcze. Tezy muszą być jednoznacznie zrozumiałe i emocjonalnie nieobojętne. Część tez jest celowo „fałszywa”, aby ograniczyć automatyzm wypowiedzi, lecz nie wolno używać zdań przeczących z partykulą *nie*. Oto przykład szczególnej tezy buforowej, jaką dla porównań warto użyć w różnych skalach służących ewaluacji edukacji:

– *Zawsze wyrażam gotowość uczenia się trudnych rzeczy.*

Pośrednie próbniki cech

Próbkowanie stanów umysłu wymaga tez sformułowanych adekwatnie do mierzonej cechy. Jednak w tezie nie powinno występować wprost słowo będące ową cechą, np. w badaniu *motywacji* błędem jest użycie formuły: – *Mam ~~motywację~~ do...*



Bodźce-stwierdzenia

Reakcje-wypowiedzi

Respondenci reagują na stwierdzenia zawarte w skali i ustosunkowują się do tez. Akceptacja lub negacja stwierdzeń jest podstawą wartościowania wskaźników i wnioskowania o polaryzacji opinii. W standaryzowanych kwestionariuszach wypowiedzi są zunifikowane, ograniczone do wyboru jednej z podanych opcji. Kolejne opcje wyznaczają narastającą wartość – od zaprzeczenia do zatwierdzenia tezy. W *Skali dwuważonych ocen* zaleca się stosowanie siedmiostopniowej *kafeterii*, np.:
absolutnie nie nie raczej nie brak zdania raczej tak tak absolutnie tak.

Pośrednie wskaźniki cech

Wybory opcji przez respondentów wskazują pośrednio, na jakich poziomach znajdują się wewnętrzne cechy i stany umysłu. W *Skali dwuważonych ocen* kateria musi składać się z symetrycznie przeciwstawnych opcji (por. dyferencjał semantyczny), z neutralną wypowiedzią pośrodku.



Reakcje-wypowiedzi

Wskaźniki intro- i ekstraspekcyjne

Ważną badawczo kwestią jest aspekt osobowy. Warto, aby respondenci wypowiadali się nie tylko o sobie, lecz także o innych podmiotach badań. Zatem introspekcyjna część wskaźników powinna dotyczyć autorefleksji, samooceny świadomości, własnych odczuć i stanów psychicznych, natomiast część ekstraspekcyjna powinna odzwierciedlać rzeczywistość, refleksję o otoczeniu i ocenę kontekstów. Pomimo subiektywności, introspekcja w ewaluacji dostarcza najwartościowsze dane empiryczne.



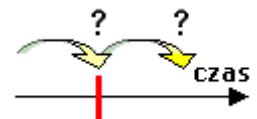
Intro- i ekstraspekcja

Przykład pary bodźców-stwierdzeń: o sobie | o innych

- *Wątpię w przydatność niniejszego opisu ukazującego istotę Skali dwuważonych ocen.*
- *Znający klasyczną metodologię badań naukowych też odniosą z tego opisu pożytek.*

Wskaźniki retro- i futurospekcyjne

Drugą kwestią jest aspekt czasowy. Respondenci powinni wypowiedzieć się o stanie aktualnym, lecz także o przewidywaniach. Retrospekcyjna część wskaźników wymaga oceny stanów wcześniejszych i procesów minionych bądź trwających, natomiast część futurospekcyjna dotyczy prognozowania stanów przyszłych i deklarowania planowanych postępowań. Jeśli stan aktualny jest niekorzystny, to perspektywa poprawy w przyszłości ma duże znaczenie wnioskotwórcze i formatywne.



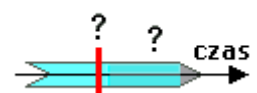
Retro- i futurospekcja

Przykład pary bodźców-stwierdzeń: dotychczas | w przyszłości

- *Dziś znam już główne zasady przygotowania i przeprowadzania Ewaluacji splotowej.*
- *Po głębszym przestudiowaniu opisu będę w stanie pojąć całą strategię Ewaluacji.*

Wskaźniki stanów krótko- i długotrwałych

Trzecią kwestią jest aspekt stałości. Respondenci winni określić, czy aktualnie zbadane stany są tylko przejściowe, czy już ustabilizowane. Warto, aby część wskaźników próbowała cechy możliwe do względnie łatwego kształtowania, a część testowała stany obiektywnie niepodatne na zmiany. W ewaluacji efektów kształcenia duże znaczenie ma trwałość formowanych cech, więc jeśli ujawnią się jedynie chwilowe trudności, to jest to mniej niepokojące niż trwała niezdolność do osiągnięcia celów.



Stany krótko- i długotrwałe

Przykład pary bodźców-stwierdzeń: chwilowo | trwale

- *Trudność ze zrozumieniem metodologii można pokonać poznając jej elementy krok po kroku.*
- *Barierą w opanowaniu sztuki ewaluacji zawsze pozostanie zawłość metod statystycznych.*

Wskaźniki cech lokalnych i globalnych

Czwartą kwestią jest aspekt zasięgu zjawiska. Chodzi o to, czy zmierzone cechy lub zjawiska dotyczą tylko grup i procesów lokalnych, czy może są typowe dla populacji. Warto, aby część wskaźników próbowała cech uczestników procesu, a część sprawdzała, jak te cechy wyglądają w społeczeństwie. W ewaluacji efektów kształcenia chodzi o badanie zjawisk lokalnych, jeśli jednak efekty są trudne do osiągnięcia, to weryfikuje się, czy może są to trudności globalne lub obiektywne.



Cechy lokalne
i globalne

Przykład pary bodźców-stwierżeń: lokalnie | ogólnie

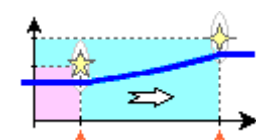
- Na studiach licencjackich jesteśmy już w stanie obliczać i oceniać podstawowe statystyki.
- Dla większości, nawet dla wielu uczonych, testy istotności statystycznej to czarna magia.

Założenia kwantytatywne

- Strategia pomiarów skalowanych jako przejście na wymiar ilościowy rezultatów badań^[1].

Próbkowanie zmienności cech

W badaniach efektów kształcenia kluczowe jest ustalenie zachodzących przemian. Oczekuje się, że nastąpią przyrosty. Jednakże w ewaluacji splotowej celowo pomija się przyrosty wiedzy, a mierzy zmiany trwalszych cech mentalnych i wolicjonalnych. Ze względu na ich trwałość, pomiędzy pomiarami zachowuje się odstęp czasu potrzebny na formowanie tych cech. Narzędzie próbkujące musi być czułe na niewielkie zmiany, choć i tak próbki psychometryczne są jedynie zgrubnym przybliżeniem stanu faktycznego.



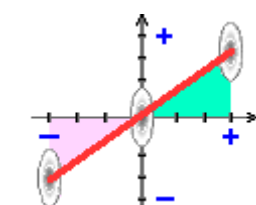
Próbkowanie zmian

Próbki na osi czasu

Za pomocą *Skali dwuważonych ocen* ustala się stany początkowe i końcowe każdej zmiennej elementarnej, u każdego respondenta. Mimo że wystąpią spadki u niektórych, to w uogólnieniu na próbę badawczą powinny nastąpić przyrosty jako wypadkowy efekt oddziaływań edukacyjnych w danym okresie.

Kwantyzacja wartości cech

W fazie kwantyzacji kolejnym progom natężenia cech jakościowych przypisuje się wagi ilościowe. Ze względu na to, że wynikające z pomiarów progi natężeń są umowne, najprościej jest przyjąć liniową zależność między stanem cech a ich wagą. W praktyce najbardziej pozytywnemu stanowi cechy nadaje się najwyższą wagę dodatnią, a skrajnie negatywnemu – wagę ujemną. W takim modelu do stanu nieokreślonego (brak zdania) przypisuje się wartość zerową. Odstępy między progami to jednostki miar, dlatego wagi są tu liczbami całkowitymi.



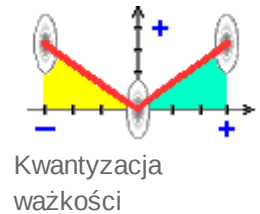
Kwantyzacja
wartości

Jakość wypowiedzi

Wyznaczona w powyższy sposób wartość cech jest ściśle zależna od jakości wypowiedzi respondentów. Jakość ta jest tym wyższa, im wyższa jest zgodność danej wypowiedzi z opinią większości co do tezy zawartej w bodźcu-stwierdzeniu.

Kwantyzacja ważkości opinii

Zastosowana do pomiarów przeciwna, „zwierciadlana” kafeteria opcji do wyboru: od negacji *absolutnie nie* do akceptacji *absolutnie tak* umożliwia analizę także drugiego wymiaru, wynikającego z intensywności wypowiedzi. Wymiar ten wskazuje, na ile ważne (znaczące dla badań) były wypowiedzi respondentów. Wyborom opcji *brak zdania* nadaje się zerową wagę, gdyż nie są przydatne poznawczo. Obu opcjom skrajnym nadaje się najwyższą wagę, choć jedna z nich jest wyborem niezgodnym z normą społeczną.



Pewność wypowiedzi

Pewność siebie w wypełnianiu ankiet zależy od osobowości. Są osoby z natury niewybierające opcji skrajnych *absolutnie*. Z kolei wybór opcji *brak zdania* może być celowym unikaniem wypowiedzi, mimo faktycznej znajomości problemu. Badanie tego wymiaru jest jednak niezbędne w pełnej, kompleksowej metodologii *Skali dwuważonych ocen*.

Założenia standaryzacyjne

- Standaryzacja jako ujednolicenie wartości liczbowych przypisywanych wskaźnikom^[1].

Rangowanie i depolaryzacja

Na potrzeby większości analiz statystycznych wystarczy nadanie wypowiedziom rang porządkowych. Prostym sposobem jest przypisanie odpowiednim opcjom wypowiedzi kolejnych liczb od 1 do 7. Pamiętać jednak należy, że część stwierżeń w skali celowo powinna być nieprawdziwa, dlatego trzeba odwrócić ich rangi w kolejności od 7 do 1. Jest to niezbędna przed zsumowaniem *depolaryzacja wskaźników*. Przyjmuje się, że zgromadzone za pomocą *Skali dwuważonych ocen* wskaźniki są addytywne, tj. upoważniają do kumulowania i do obliczania wartości średnich.

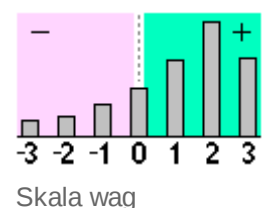


Racja większości

Polaryzacja danego *Wskaźnika* (znak walencyjny + lub -) wynika z prawdziwości bądź fałszywości bodźca-stwierdzenia. Musi to być jednak empirycznie potwierdzone przez kwalifikowaną większość respondentów. Jeśli większość ma inne zdanie niż twórca narzędzia, to trzeba takie niepewne stwierdzenie usunąć ze skali.

Kalibracja w skali wag

W ścisłej metodologii *Skali dwuważonych ocen* wypowiedziom nadaje się wagi od -3 do +3. Umożliwia to analizę statystyczną obu opisanych powyżej wymiarów, gdyż wartość cech oblicza się jako średnią arytmetyczną wag, a ważkość opinii jako średnią z wartości bezwzględnych tych wag. Dzięki depolaryzacji, w prawidłowo zbudowanych skalach uogólnione poziomy cech po uśrednieniu dla grup powinny osiągać zawsze wartości dodatnie. Obszar rezultatów od 0 do +3 nazywany jest *strefą zgodności*.

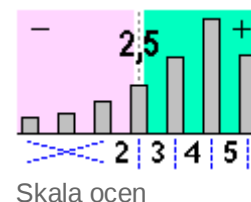


Zakresy zmiennych

W publikowaniu wyników badań konieczne jest podanie zakresu wag lub rang zastosowanych do kalibracji skali pomiarowej. W skali wag można łatwo wystandaryzować poziomy cech tak, aby uzyskać maksimum = 1, dzieląc rezultaty średnie przez 3.

Konwersja na skalę ocen

Dobłą do publikacji reprezentacją rezultatów jest wyrażenie ich w skali ocen akademickich od 2 do 5. Ma to tę zaletę, że w polskim społeczeństwie jest silnie zakorzeniony skrypt łatwej interpretacji ocen. Trafnie wyczuwamy różnice w jakości cech, które zmieniają poziomy np. z 3,85 na 4,5 lub z 4,0 na 4,35. Skala akademicka jest tu lepsza niż szkolna, ze względu na spójne nazwy poziomów jakościowych (*niedostateczny, dostateczny, dobry, bardzo dobry*). Konwersja wyników na skalę ocen jest prosta – wystarczy dodać wartość 2,5 do obliczonych w skali wag miar wyznaczających poziomy (np. średnie). Natomiast w konwersji ze skali rang należy odjąć wartość 1,5.



Przedziały ocen

– Z czego to wynika? Wybór opcji *brak zdania* nie jest ani akceptacją, ani negacją. Na studiach odpowiada to progowi między *ndst* a *dst*, gdy jest dylemat: zaliczyć czy nie, stąd wartość progowa 2,5. Z kolei wartość maksymalna 5,5 ma uzasadnienie w tym, że wokół ocen głównych występują odchylenia $\pm$ pół działki, co skutkuje podziałem skali także na przedziały o progach 3,5 i 4,5. Wartości cech poniżej 2 i powyżej 5 po uśrednieniu dla prób w praktyce badań nie występują.

Założenia metodologiczne i analityczne

PATRZ Ewaluacja splotowa i Miary dwuważonych ocen

Zobacz też

- Ewaluacja
- Skala Likerta

Przypisy

1. Stanisław Ubermanowicz: *Ewaluacja splotowa InfoKultury. Skala dwuważonych ocen* (<http://repozytorium.amu.edu.pl/bitstream/10593/946/3/EwaluacjaSplotowa.pdf>), Poznań: Wydawnictwo Naukowe UAM, 2005.
2. Elżbieta Hornowska: *Operacjonalizacja wielkości psychologicznych. Założenia – struktura – konsekwencje*. Wrocław: Zakład Narodowy im. Ossolińskich, 1989.

Bibliografia

- Brzeziński Jerzy: *Metodologia badań psychologicznych*. Warszawa: Wydawnictwo Naukowe PWN, 2004. ISBN 83-01-14041-0.
- Hornowska Elżbieta: *Operacjonalizacja wielkości psychologicznych. Założenia – struktura – konsekwencje*. Wrocław: Zakład Narodowy im. Ossolińskich, 1989. ISBN 83-04-03309-7.

- Ubermanowicz Stanisław: *Ewaluacja splotowa InfoKultury. Skala dwuważonych ocen* (<http://repozytorium.amu.edu.pl/bitstream/10593/946/3/EwaluacjaSplotowa.pdf>). Poznań: Wydawnictwo Naukowe UAM, 2005. ISBN 83-232-1608-8.
-

Źródło: „https://pl.wikipedia.org/w/index.php?title=Skala_dwuważonych_ocen&oldid=63771507”

Tę stronę ostatnio edytowano 21 cze 2021, 10:11. Tekst udostępniany na licencji Creative Commons: uznanie autorstwa, na tych samych warunkach (<http://creativecommons.org/licenses/by-sa/3.0/deed.pl>), z możliwością obowiązywania dodatkowych ograniczeń. Zobacz szczegółowe informacje o warunkach korzystania (http://foundation.wikimedia.org/wiki/Warunki_korzystania).