

dr Monika Jakubowska-Mirek

Uniwersytet Warszawski

dr Ewa Stożek

Pewność wyboru odpowiedzi w zadaniach zamkniętych

W ramach szerokiej współpracy Wydziału Pedagogicznego Uniwersytetu Warszawskiego oraz Urzędu Miasta Ostrołęka przeprowadzana jest, między innymi, diagnoza osiągnięć uczniów z matematyki na rok przed egzaminem ósmoklasisty. Uczniowie rozwiązują test, który powinien dać odpowiedź, w jakim stopniu są oni gotowi do osiągnięcia sukcesu na egzaminie ósmoklasisty oraz jakie braki wymagają dalszej pracy. Tegoroczne badanie było przeprowadzone 9 czerwca 2022 roku na grupie 586 uczniów klas siódmych szkół podstawowych Ostrołęki, włączając w to uczniów ze specjalnymi potrzebami edukacyjnymi oraz grupę uczniów – obywateli Ukrainy. Narzędzie diagnostyczne składało się z 20 zadań: 10 zadań zamkniętych wielokrotnej odpowiedzi (należało wybrać jedną poprawną odpowiedź: A, B, C lub D) oraz 10 zadań otwartych. Za cały test uczeń mógł maksymalnie uzyskać 30 punktów.

W egzaminie ósmoklasisty z matematyki liczba zadań zamkniętych znacząco przewyższa liczbę zadań otwartych, a co za tym idzie, umiejętność stosowania różnorodnych strategii rozwiązywania zadań zamkniętych przez ucznia jest ważna dla jego wyniku. Dla badaczy-diagnostów ma znaczenie, czy odpowiedzi zaznaczone są przez uczniów losowo, czy też stoi za tym świadomy wybór, związany z jakąś strategią (eliminację mało prawdopodobnych odpowiedzi, wybór najbardziej prawdopodobnej odpowiedzi, sprawdzenie poprawności każdej z odpowiedzi czy wreszcie „otwieranie” zadania zamkniętego). Nie mamy wglądu w to, jak uczniowie rozwiązują zadania zamknięte (inaczej niż w przypadku zadań otwartych), ale możemy ich zapytać, jak bardzo pewni są wyboru swojej odpowiedzi.

W tegorocznym badaniu umiejętności matematycznych w klasie VII poprosiliśmy uczniów, by określili, jak pewni są odpowiedzi, którą zaznaczyli w zadaniach zamkniętych. Uczniowie otrzymali następującą instrukcję:

W zadaniach od 1. do 10. zaznacz odpowiedni emotikon:

- ☺ – jestem pewny(-na), że to poprawna odpowiedź,
- ☹ – mam wątpliwości, czy wybrałem(-łam) poprawną odpowiedź,
- ☹ – odpowiedź wybrałem(-łam) na chybił trafił.

Przykład zaznaczenia wskazującego, że wybrałeś/wybrałaś odpowiedź na chybił trafił.

Zadanie 1. (0–1)



Przyznajemy, że wprowadzając emotikony do naszego testu, nie liczyliśmy na wiele. Nasze oczekiwania dotyczyły zagadnień dydaktycznych (dodatkowe informacje dla nauczycieli, które mogą uatrakcyjnić coroczne podsumowanie diagnozy). Uznałyśmy, że informacje na temat pewności uczniowskich odpowiedzi mogą stanowić punkt wyjścia do rozmowy z nauczycielami o strategii rozwiązywania zadań zamkniętych. Planowałyśmy wskazać typy zadań, w których uczniowie najczęściej „strzelają”. Jednak gdy zaczęłyśmy eksplorować zgromadzone dane, okazało się, że jest w nich coś więcej. W naszym referacie chcielibyśmy podzielić się pierwszymi spostrzeżeniami po wstępnej analizie zgromadzonych danych.

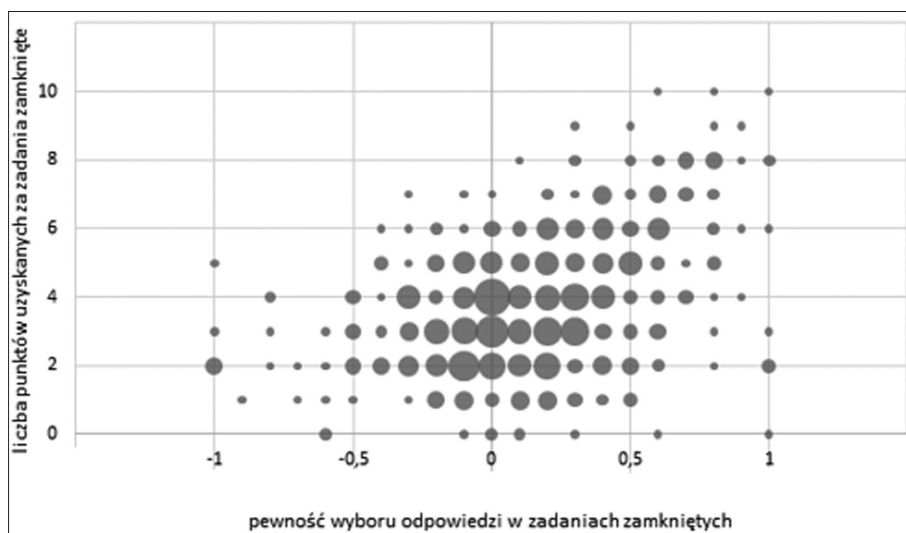
Co widać w danych?

Zaprezentowane tu obserwacje ograniczyliśmy do grupy osób, które na egzaminie ósmoklasisty rozwiązywałyby zadania arkusza standardowego, są to 572 osoby, w tym 304 dziewcząt (53%) i 268 chłopców (47%) i jednocześnie przy każdym zadaniu zamkniętym jednoznacznie wybrały emotikon. Ostatecznie analizowana tutaj grupa to 465 uczniów, w tym 243 dziewcząt (52%) i 222 chłopców (48%).

To, co nas najbardziej niepokoiło przed badaniem, to wiarygodność pozyskiwanych w ten sposób informacji. Czy uczniowie będą zaznaczać emotikony? Czy podejną do tego poważnie czy raczej „rozrywkowo”? Czy będą mieli odwagę się przyznać, że wybierali odpowiedź na chybił trafił? Nasze obawy nie potwierdziły się. Uczniowie chętnie zaznaczali emotikony, niektórzy nawet dorysowywali je przy zadaniach otwartych, by poinformować, jak bardzo pewni są swojego rozwiązania. Nie mieli oporów, by przyznać się do tego, że zaznaczyli odpowiedź na chybił trafił (w zależności od zadania od 4% do 45% uczniów), przy czym liczba „trafień” (przy wyborze L) dla większości zadań znajduje się w przedziale od 20% do 30%, co wskazywałoby faktycznie na ich losowy charakter (w idealnej sytuacji spodziewałibyśmy się 25% trafień).

Emotikonom zostały przypisane wartości liczbowe $J = 1$, $K = 0$, $L = -1$. Jako syntetyczny wskaźnik p – pewności wyboru (dokładniej: poczucia pewności wyboru) odpowiedzi w zadaniach zamkniętych przyjęliśmy średnią wartość „emotikonów” ze wszystkich, dziesięciu zadań zamkniętych. Wskaźnik p przyjmuje wartości od -1 do 1 ($p = -1$ oznacza, że uczeń wybierał odpowiedź na chybił trafił we wszystkich zadaniach zamkniętych, $p = 1$ oznacza, że uczeń we wszystkich zadaniach zamkniętych był pewny swojej odpowiedzi; ujemna wartość wskaźnika p oznacza, że uczeń częściej strzelał, niż był pewny odpowiedzi, dodatnia wartość – częściej był pewny odpowiedzi niż strzelał na chybił trafił). Następnie te wyniki zestawiliśmy z liczbą poprawnych odpowiedzi za zadania zamknięte (rys. 1).

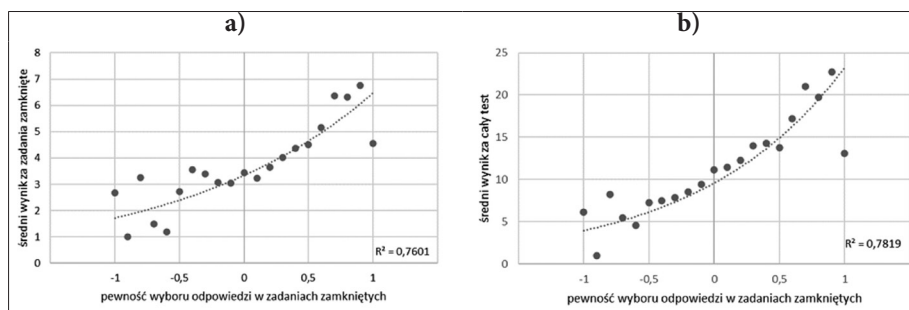
Interesowała nas zależność uczniowskiej pewności wyboru i poprawności tego wyboru. Okazuje się, że spektrum możliwości dotyka obu biegunów: wśród badanych znalazł się uczeń, który przy pewności 1 (był pewny swoich odpowiedzi we wszystkich zadaniach zamkniętych) uzyskał maksymalną liczbę punktów, ale też taki, który przy pewności 1 nie uzyskał żadnego punktu za zadania zamknięte.



Rysunek 1. Zależność między pewnością wyboru odpowiedzi w zadaniach zamkniętych a liczbą punktów za zadania zamknięte, (N = 465)

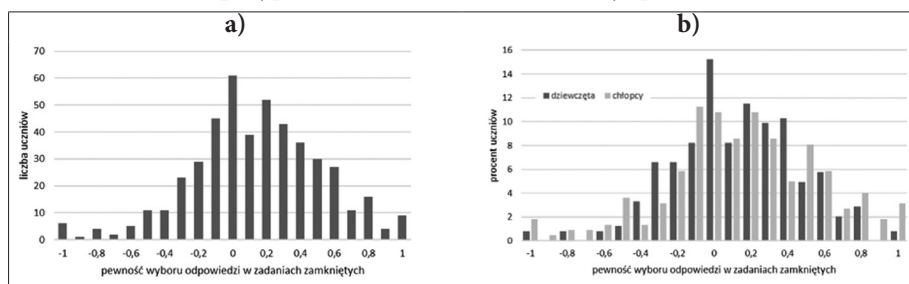
Uczeń mógł maksymalnie uzyskać 10 punktów za zadania zamknięte. Wielkość „bąbla” odpowiada liczbie uczniów, którzy uzyskali ten sam wynik za zadania zamknięte przy tym samym poziomie pewności wyboru odpowiedzi w zadaniach zamkniętych. Dla przykładu, największy „bąbel” odpowiadający pewności = 0 oraz liczbie punktów = 4 obrazuje wyniki dla 19 uczniów.

Na podstawie wyników widać wyraźny związek między pewnością odpowiedzi w zadaniach zamkniętych a poziomem umiejętności matematycznych, mierzonym wynikami testu. Głębsza analiza tej zależności wymaga oczyszczenia danych (między innymi identyfikacji uczniów o lekceważącym stosunku do całego testu), żeby określić wartość losowych wyników oraz dodatkowych parametrów, by lepiej skalibrować dane. Jednak już na etapie wstępnej analizy można uznać, że istnieje zależność pewności i trafności odpowiedzi uczniowskich. Tę zależność pokazuje rysunek 2a, na którym zestawiono pewność wyboru odpowiedzi w zadaniach zamkniętych ze średnim wynikiem za zadania zamknięte. Jeśli taki związek istnieje, to znaczy, że pewność wyboru odpowiedzi w zadaniach zamkniętych dobrze opisuje poziom wiedzy i umiejętności ucznia. A jeśli tak, to można się spodziewać, że cecha pewności wyboru odpowiedzi w zadaniach zamkniętych powinna być powiązana z ogólnym wynikiem testu (przypomnijmy, że za zadania otwarte uczeń mógł uzyskać dwa razy więcej punktów niż za zadania zamknięte). Zależność tę pokazano na rysunku 2b. W obu przypadkach (zadań zamkniętych oraz całego testu) zależność jest wykładnicza.



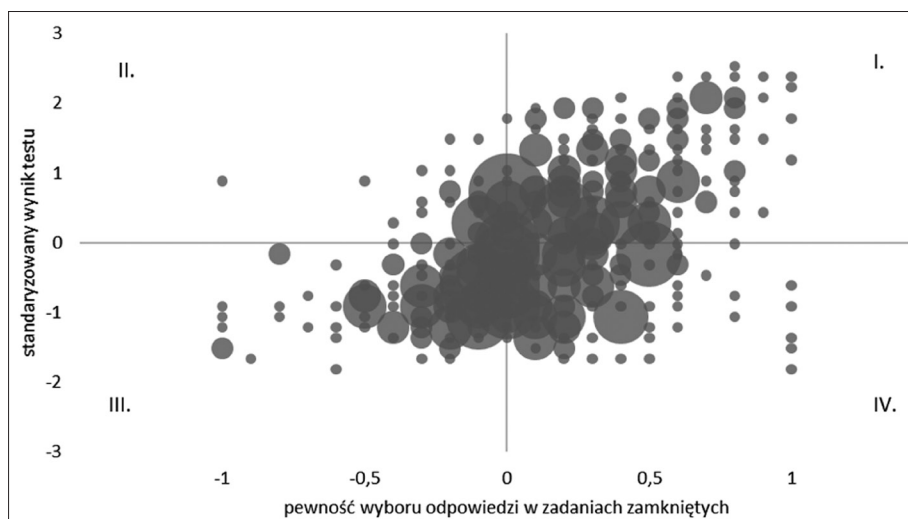
Rysunek 2. Zależność między pewnością wyboru odpowiedzi w zadaniach zamkniętych i a) średnim wynikiem za zadania zamknięte, b) średnim wynikiem za cały test, (N = 465)

Rozkład cechy pewność wyboru odpowiedzi w zadaniach zamkniętych (rys. 3) jest zbliżony do normalnego, ze średnią 0,14 i odchyleniem standardowym 0,38. Nie widać też znaczących różnic w rozkładzie dla dziewcząt i chłopców (średnia dla dziewcząt 0,13, odchylenie standardowe 0,35; średnia dla chłopców 0,15 i odchylenie standardowe 0,42). Rozkład pewności wyboru odpowiedzi w zadaniach zamkniętych pokazuje, że dziewczęta ogólnie nie są mniej pewne niż chłopcy, choć częściej niż chłopcy przyjmują ostrożną postawę (może tak, może nie; mam wątpliwości $p = 0$), natomiast chłopcy częściej są bardzo pewni siebie (co w wielu przypadkach oznacza również: zbyt pewni siebie).



Rysunek 3. Rozkład cechy pewność wyboru odpowiedzi w zadaniach zamkniętych a) dla wszystkich uczniów, N = 465 oraz b) z podziałem na płeć - 243 dziewcząt i 222 chłopców

Popatrzmy zatem na zależność między pewnością wyboru odpowiedzi w zadaniach zamkniętych a ogólnym wynikiem testu z punktu widzenia indywidualnych wyników, a nie tylko wyników uśrednionych (rys. 4).



Rysunek 4. Zależność między pewnością wyboru odpowiedzi w zadaniach zamkniętych a wynikiem ogólnym testu (N = 465)

Średnia testu 12,1 punktu (40% możliwych do uzyskania punktów), odchylenie standardowe 6,7 punktu. Wielkość „bąbla” odpowiada liczbie uczniów, którzy uzyskali ten sam wynik testu przy tym samym poziomie pewności wyboru odpowiedzi w zadaniach zamkniętych.

Uczniów, ze względu na pewność wyboru odpowiedzi w zadaniach zamkniętych oraz osiągnięte wyniki z całego testu, można podzielić na cztery grupy, opisane poniżej. Do charakterystyk każdej grupy dołączyliśmy średnią ocen z matematyki z I semestru klasy VII, która może posłużyć jako kryterium trafności oceny wiedzy i umiejętności uczniów.

Grupa I. Uczniowie, którzy są pewni swoich wyborów w zadaniach zamkniętych ($p > 0$) i osiągający wyniki powyżej średniej. W tym przypadku poczucie pewności wynika z rzeczywistej wiedzy i umiejętności uczniów. Tę grupę nazwaaliśmy „samoświadomi” – sądzą, że wiedzą i faktycznie wiedzą.

Grupa I	Liczba uczniów	Średni standaryzowany wynik za cały test	Średnia pewność	Średnia ocen z matematyki (I semestr klasy VII)
Ogółem	158	1,06	0,44	4,39
Dziewczęta	83	0,97	0,41	4,34
Chłopcy	75	1,16	0,46	4,44

Grupa II. Uczniowie, którzy nie są pewni swoich wyborów w zadaniach zamkniętych ($p \leq 0$), ale mimo to uzyskują wyniki powyżej średniej. To najmniej liczna grupa. W tej grupie dziewcząt jest dwukrotnie więcej niż chłopców. Nazwaaliśmy tę grupę „asekuranci” – wątpią, że wiedzą, a jednak wiedzą.

Grupa II	Liczba uczniów	Średni standaryzowany wynik za cały test	Średnia pewność	Średnia ocen z matematyki (I semestr klasy VII)
Ogółem	44	0,62	-0,11	4,09
Dziewczęta	30	0,67	-0,10	4,17
Chłopcy	14	0,51	-0,14	3,93

Grupa III. Uczniowie, którzy nie są pewni swoich wyborów ($p \leq 0$) i osiągają wyniki średnie i poniżej średniej. Nazwaaliśmy tę grupę uczniów „przegranii” – wątpią, że wiedzą i faktycznie nie wiedzą.

Grupa III	Liczba uczniów	Średni standaryzowany wynik za cały test	Średnia pewność	Średnia ocen z matematyki (I semestr klasy VII)
Ogółem	154	-0,76	-0,24	2,73
Dziewczęta	76	-0,69	-0,22	2,82
Chłopcy	78	-0,83	-0,26	2,64

Grupa IV. Uczniowie, którzy są pewni swoich wyborów ($p > 0$), ale osiągają wyniki na poziomie średnim i poniżej średniego. To grupa uczniów nieadekwatnie oceniających swoją wiedzę i umiejętności z matematyki, są zbyt pewni swoich możliwości, które nie znajdują potwierdzenia w wynikach testu. Tę grupę uczniów nazwaaliśmy „celebryci” – sądzą, że wiedzą, a jednak nie wiedzą.

Grupa IV	Liczba uczniów	Średni standaryzowany wynik za cały test	Średnia pewność	Średnia ocen z matematyki (I semestr klasy VII)
Ogółem	109	-0,72	0,34	2,96
Dziewczęta	54	-0,68	0,30	2,96
Chłopcy	55	-0,75	0,38	2,96

Wstępny ogląd zebranych danych pozwala na sformułowanie następujących wniosków:

1. Zgromadzone dane dotyczące poczucia pewności przy wyborze odpowiedzi w pytaniach zamkniętych są wiarygodne, a nastawienie uczniów adekwatne do celów badania (uczniowie nie zignorowali prośby dotyczącej zaznaczenia odpowiedniego emotikonu i wykonali to polecenie w sposób odpowiedzialny).
2. Uczniowskie poczucie pewności wyboru odpowiedzi w zadaniach zamkniętych jest powiązane nie tylko z wynikiem za zadania zamknięte, ale także z ogólnym wynikiem testu i może mieć związek z doświadczeniami (pozytywnymi lub negatywnymi), a co za tym idzie postawą uczniów w odniesieniu do matematyki (np. „przegranii” to uczniowie, których charakteryzuje bezradność matematyczna i często brak motywacji do rozwiązywania zadań. Ci uczniowie uwierzyli, że nie umieją matematyki). Taka wiedza jest bardzo ważna dla nauczyciela, który

na jej podstawie może konstruować odpowiednią informację zwrotną i korygować nie tylko wiedzę i umiejętności, ale także postawę ucznia. Wpływ przekonań dotyczących własnych kompetencji matematycznych na osiągnięcia matematyczne został potwierdzony przez wielu badaczy (m.in. Pinxten, 2013; Ganley i Lubieński, 2016).

3. Poczucie pewności wyboru odpowiedzi dziewcząt i chłopców w badanej populacji nie różni się w sposób znaczący. Należy jednak zauważyć, że dziewczęta wykazały się większą niepewnością w odniesieniu do swoich odpowiedzi. Wśród uczniów, którzy na pytania zamknięte odpowiedzieli w sposób poprawny i nie byli pewni swoich odpowiedzi, udział procentowy dziewcząt jest znacznie większy. Co ciekawe, „asekurantki” mają też lepsze oceny niż „asekuranci”, a zatem potencjalnie większy powód, by bardziej ufać swoim odpowiedziom.
4. Na podstawie uzyskanych wyników można wyodrębnić cztery grupy uczniów, różniące się zależnością pewności wyboru odpowiedzi w pytaniach zamkniętych od wyników ogólnych testu. Wstępna charakterystyka tych grup może być punktem wyjścia do rozmów z nauczycielami na temat wsparcia uczniów w nauczaniu matematyki. Kluczowe wydaje się w tym przypadku adekwatne formułowanie komunikatów zwrotnych do uczniów – innych do tych, którzy nie ufają swojej wiedzy matematycznej, mimo skutecznych strategii rozwiązań, i innych do uczniów, którzy są zbyt pewni siebie i nie zauważają popełnianych przez siebie błędów.

Warto podkreślić, że zastosowane przez nas narzędzie jest z praktycznego punktu widzenia bardzo atrakcyjne. Zostało ono wykorzystane „przy okazji”, a jego przygotowanie nie wymagało dużego nakładu pracy. Tymczasem zebrane dane wydają się atrakcyjne, a pierwsze ich analizy obiecujące. Te dane inspirują do wielu analiz, zarówno dotyczących konstrukcji zadań testowych (podatność na zgadywanie, funkcjonowanie zadania w teście, diagnostyczna wartość zadania), jak i tych dotyczących samych uczniów i ich samooceny. Pozwalają lepiej zrozumieć uwarunkowania odpowiedzi uczniów związane z ich postawami.

Czy próbowano w podobny sposób zbierać informacje o pewności uczniowskich wyborów w zadaniach zamkniętych?

Oczywiście jest to pytanie, które powinnyśmy sobie postawić przed przystąpieniem do zbierania danych, ale jak napisałyśmy we wstępie, pierwotnie myślałyśmy o tych danych jako o punkcie wyjścia do rozmowy z nauczycielami matematyki o strategii rozwiązywania zadań zamkniętych.

Okazuje się, że prowadzone są duże badania, w których takie dane są zbierane (Foster i in., 2022). Wykorzystują platformę on-line *Eedi* z zamkniętymi diagnostycznymi zadaniami matematycznymi do zbierania danych o poczuciu pewności wyboru odpowiedzi. Badaniem objęto $N = 7302$ uczniów w wieku od 6 do 16 lat, którzy w sumie udzielili 219 826 odpowiedzi, co oznacza, że średnio każdy uczeń rozwiązał 30 zadań na tej platformie. W badaniu wykorzystano 5-stopniową skalę Likerta i też zaimplementowano ją w formie emotikonów.

Badacze, którzy stosowali podobne narzędzia w testach matematycznych, podkreślali między innymi ich wartość dydaktyczną. Wiedza na temat uczniowskiego poczucia pewności przy wyborze odpowiedzi na pytania zamknięte jest bardzo ważna dla nauczycieli, którzy opierają na niej informację zwrotną dla ucznia (Foster, 2016).

Często w tego typu badaniach wykorzystywane są dodatkowe narzędzia kwestionariuszowe, takie jak Skale Pewności w Uczeniu się Matematyki (*Confidence in Learning Mathematics Scale*) czy Skale Postaw Matematycznych (*Mathematics Attitudes Scales*) (Fennema i Sherman, 1976; Lim i Chapman, 2013). Wyniki analiz wskazują silną korelację pewności odpowiedzi z samooceną dotyczącą umiejętności matematycznych. Pewność siebie uczniów jest powiązana z ich motywacją i wyborami (Bandura, 1997; Bouffard i Narciss, 2011).

Wielu badaczy wykazało, że dziewczęta wykazują mniejszą pewność siebie podczas wyboru odpowiedzi w pytaniach zamkniętych, a także że zasadniczo w mniejszym stopniu ufają swoim umiejętnościom matematycznym (Ganley i Lubiński, 2016; Mullis i in., 2020). Wyniki badania Foster i in. (2021) wskazują na wyższą średnią pewność wyboru u chłopców i jednocześnie brak istotnych różnic między chłopcami i dziewczętami w odniesieniu do uzyskanych wyników. Badacze zauważyli, że nadmierna pewność poprawności wyboru przez chłopców jest bardziej widoczna w przypadku uczniów osiągających gorsze wyniki (Foster i in., 2022).

Podziękowania

Dziękujemy Joannie Kowalskiej (OKE w Łodzi)
za inspirację do badania pewności wyboru odpowiedzi w zadaniach zamkniętych.

Bibliografia

- Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*, Freeman, New York.
- Bouffard, T., Narciss S. (2011). Benefits and risks of positive biases in self-evaluation of academic competence. *Introduction International Journal of Educational Research*, 50 (4), 205–208. <https://doi.org/10.1016/j.ijer.2013.10.008>
- Fennema, E., Sherman, J. A. (1976). Fennema-Sherman mathematics attitudes scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education*, 7(5), 324–326. <https://doi.org/10.2307/748467>
- Foster, C. (2016). Confidence and competence with mathematical procedures, *Educ Stud Math*, 91: 271–288. <https://doi.org/10.1007/s10649-015-9660-9>
- Foster, C., Woodhead, S., Barton I C., Clark-Wilson, A. (2022). School students' confidence when answering diagnostic questions online, *Educational Studies in Mathematics* (2022) 109: 491–521. <https://doi.org/10.1007/s10649-021-10084-7>
- Ganley, C. M., Lubiński, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. *Learning and Individual Differences*, 47, 182–193. <https://doi.org/10.1016/j.lindif.2016.01.002>

- Lim, S. Y., Chapman, E. (2013). Development of a short form of the attitudes toward mathematics inventory. *Educational Studies in Mathematics*, 82(1), 145–164. <https://doi.org/10.1007/s10649-012-9414-x>
- Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/international-results/>
- Pinxten, M., Marsh, H. W., De Fraine, B., Van Den Noortgate, W., Van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *British Journal of Educational Psychology*, 84(1), 152–174. <https://doi.org/10.1111/bjep.12028>