

JAK TO JEST Z TYMI OCENAMI (a raczej
ocenianiem) W SZKOLE?

CZEGO UCZĄ NAS WYNIKI BADAŃ
I DOŚWIADCZENIA INNYCH KRAJÓW?

Jacek Pyżalski

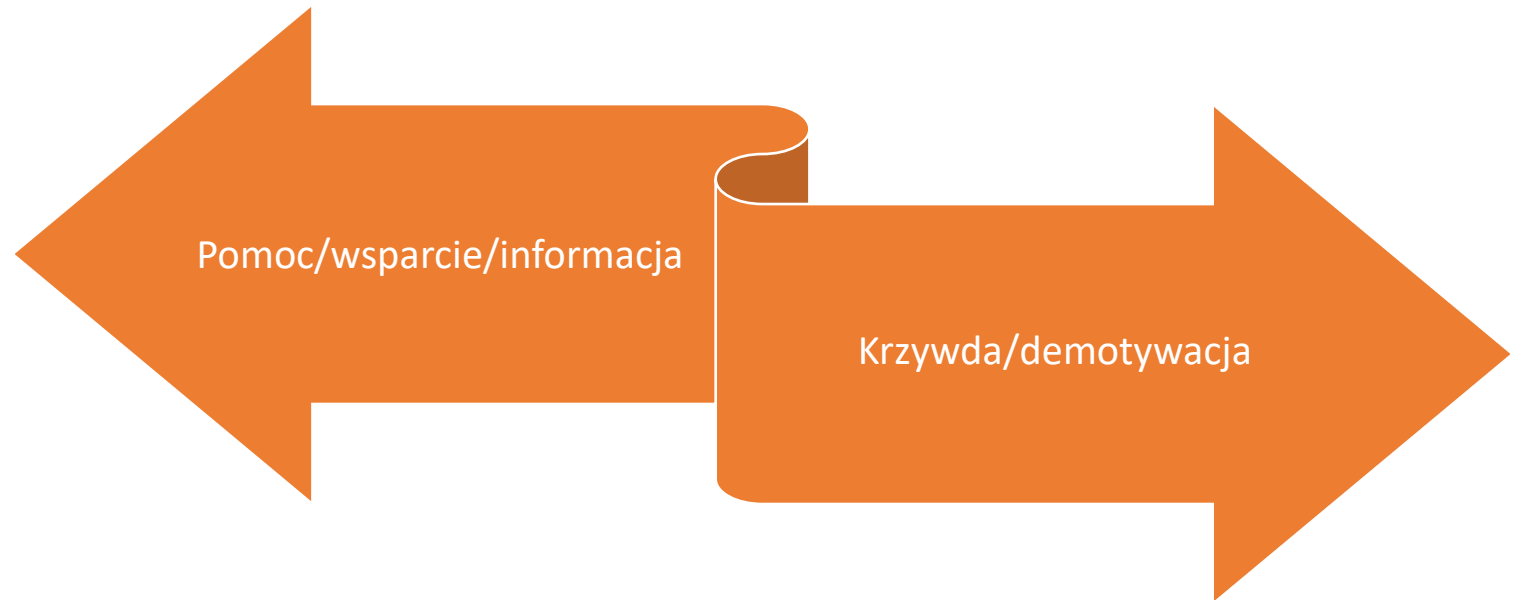
Wydział Studiów Edukacyjnych

Uniwersytet im. A. Mickiewicza w Poznaniu



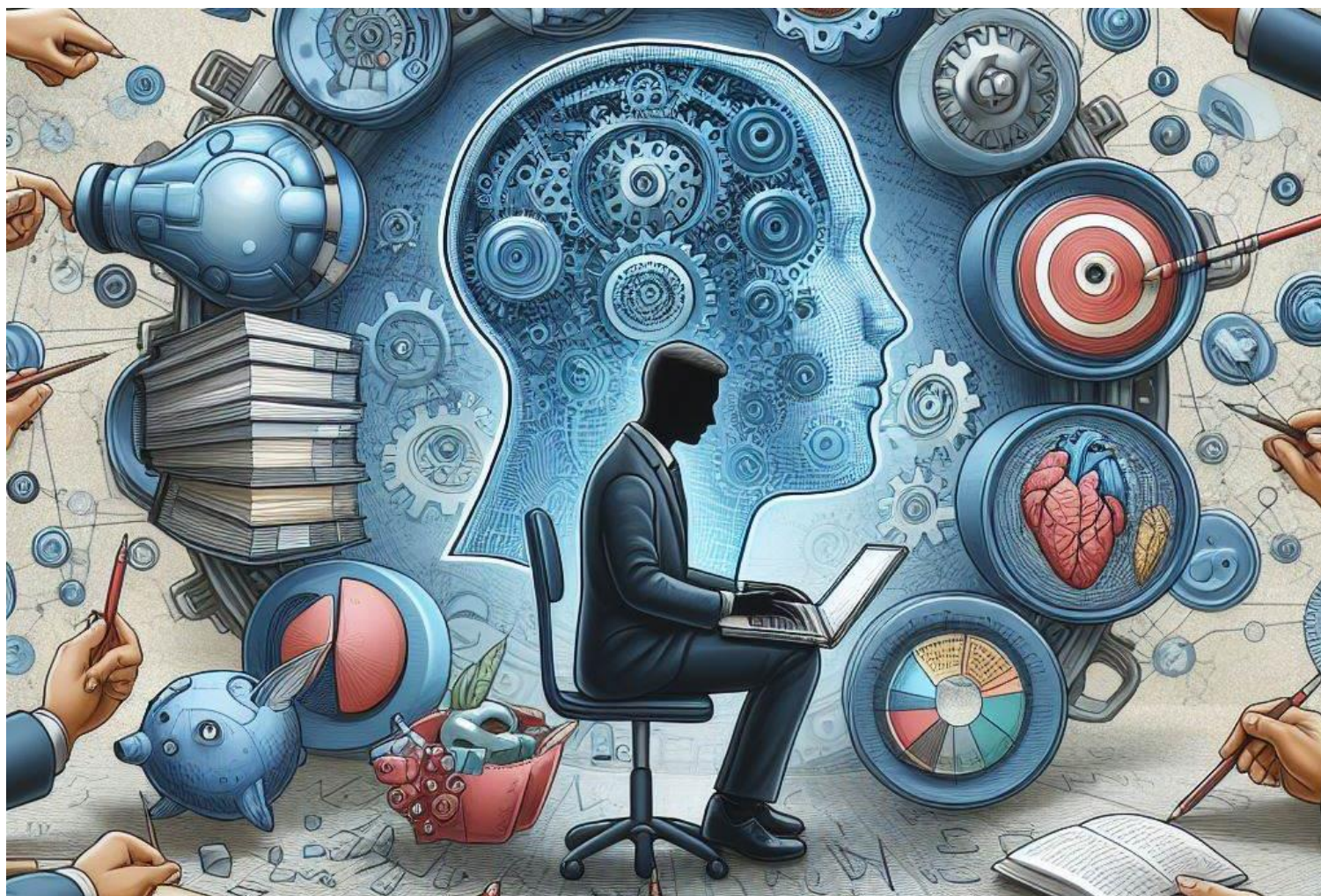
Dlaczego się temu
przyjrzałem?

Ocenianie
nie jest
neutralne i
budzi
emocje





Jedna z wielu rzeczy postrzeganych jako balast w edukacji,
który po wyrzuceniu spowoduje, że uniesiemy się w górę



Duże
znaczenie,
wiele
wymiarów
(także
ukrytych)

Analiza tego tematu nie jest łatwa...

- Sporo badań, w różnych kontekstach
- Ale jednocześnie wiele kwestii niedobadanych
- Pomieszanie porządku empirycznego z normatywnym
- Rozwiązania legislacyjne – jakość jest
- Tradycja
- Temperatura dyskusji publicznych nt. ocen (!)



Nie da się pominąć wymiaru
etycznego oceny!

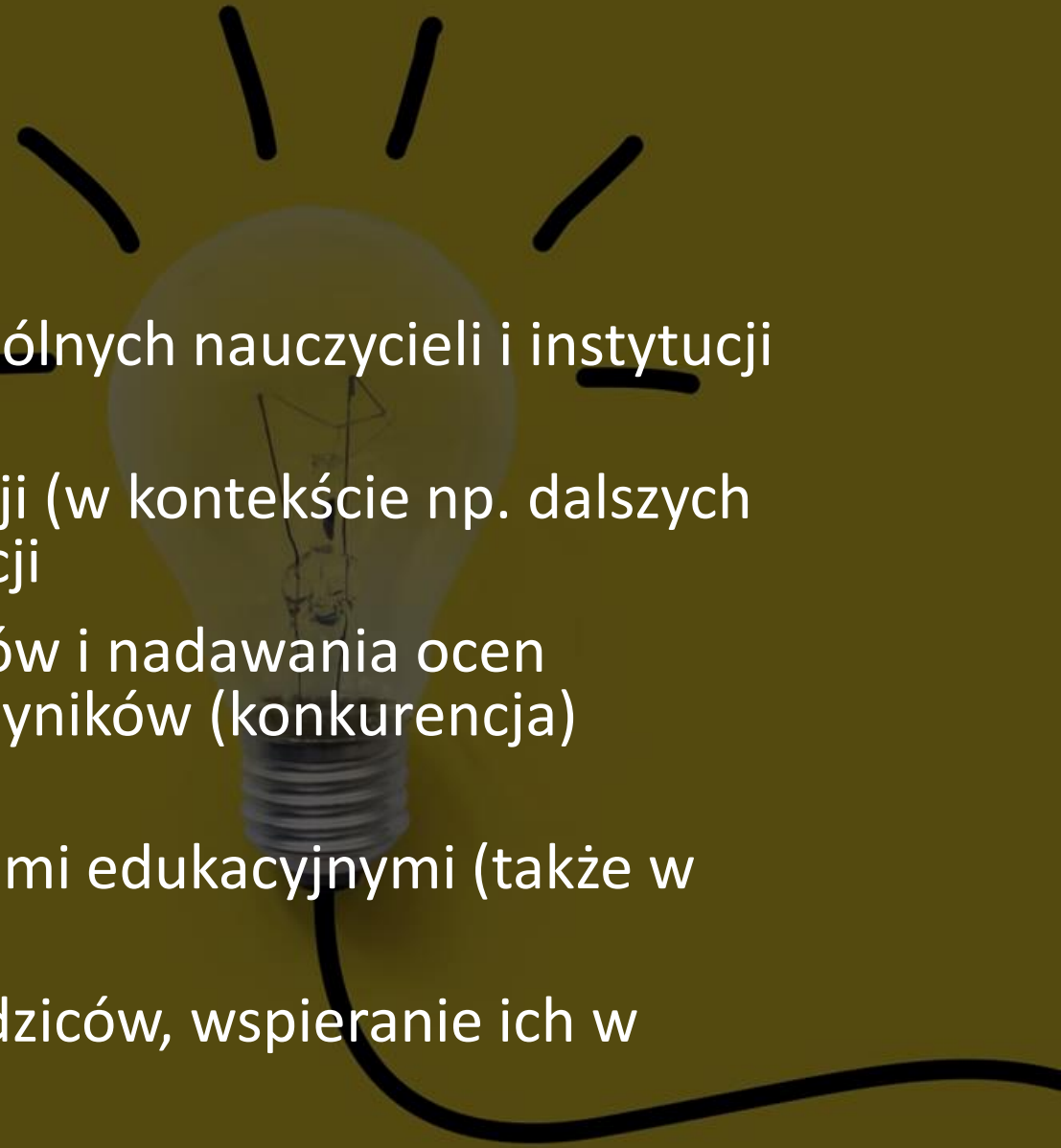
Ocena

W obszarze dydaktyki

W obszarze
zachowania/postępowania

Po co w ogóle oceniamy?

- ocena jakości i efektywności pracy poszczególnych nauczycieli i instytucji edukacyjnych
- identyfikacja możliwości osób w celu selekcji (w kontekście np. dalszych niedostępnych dla wszystkich etapów edukacji)
- proces oceniania efektów uczenia się uczniów i nadawania ocen poszczególnym osobom na podstawie tych wyników (konkurencja) (Heywood, 1989)
- standaryzacja komunikacji między instytucjami edukacyjnymi (także w aspekcie międzynarodowym)
- informacja zwrotna dla uczących się i ich rodziców, wspieranie ich w procesie



Przykład

A= 90-100%

B= 80-89%

C= 70-79%

D= 60-69%

F= 0-59%

Ale czy zawsze tak jest?

Czy oceny mają coś wspólnego z sukcesem w życiu?

- Mają – zarobki, kariera zawodowa (ale siła związku jest raczej przeciętna) (Afarian i in., 2003, Cohen, 1984 – metaanaliza)
- Pytanie: co jest sukcesem w życiu?

Afarian, R., & Kleiner, B. H. (2003). The relationship between grades and career success. *Management Research News*, 26(2/3/4), 42-51.

Cohen, P. A. (1984). College grades and adult achievement: A research synthesis. *Research in Higher Education*, 20, 281-293.

Podsumowanie, które daje do myślenia...

Wyniki przedstawione w niniejszym artykule mogą być nieco zniechęcające dla tych, którzy przywiązują dużą wagę do ocen i ich wartości predykcyjnej. Wydaje się, że to, jak dobrze uczeń radzi sobie na studiach, ma tylko marginalny związek z sukcesem w karierze. Jak zauważył McClelland (1973), co ważniejsze dla sukcesu zawodowego jest to, czy ktoś w ogóle uczęszcza na studia. Należy jednak zauważyć, że główny nacisk w tej syntezie został położony na osiągnięcia zawodowe, co stanowi dość ograniczoną definicję sukcesu dorosłego życia. Niewiele badań dotyczyło związków z innymi kryteriami sukcesu życiowego, takimi jak szczęście małżeńskie i rodzinne, zaangażowanie w społeczność, itp. Mając to na uwadze, badacze mogą rozważyć poszerzenie swoich definicji sukcesu dorosłego w przyszłości (Cohen, 1984)



Najbardziej znienawidzone pytanie...

Czy to na ocenę?

Dlaczego nie lubimy tego pytania?



Liczą się szczegóły

Deci, Koestner i Ryan (2001) w swoich bardzo interesujących badaniach bliżej skoncentrowali się na szczegółowych kwestiach. Prowadzone przez nich analizy wyników wielu badań sugerują, że **wzmocnienia werbalne wzmacniają motywację wewnętrzną**. Nie mogą one mieć jednak charakteru kontrolującego, gdy na przykład nauczyciel chwali uczniów za to, że poddali się jego poleceniom. Inne nagrody nie mają już takiego wpływu na motywację wewnętrzną, z jednym wszakże wyjątkiem. **Nagrody materialne i symboliczne, na przykład odznaka, przekładały się na motywację wewnętrzną negatywnie.**



Czy słowo się liczy – ale jakie i do kogo?

- Liczą się bardzo – to jest to, co się pamięta często po latach – nie muszą być rozbudowane (tu też jest negatywna strona)
- Eksperyment – co i jak mówimy, jak ktoś zada ciekawe pytanie, albo coś interesująco przemyśli.
- Czy pochwała jest manipulacją?



Jest jedno ciekawe badanie...

Consequences of school grading systems on adolescent health: evidence from a Swedish school reform

Björn Hagberg, Joakim Lindgren, Klara Johansson, Mattias Strandh & Solveig Petersen Björn Högberg, Joakim Lindgren, Klara Johansson, Mattias Strandh & Solveig Petersen (2021) Consequences of school grading systems on adolescent health: evidence from a Swedish school reform, *Journal of Education Policy*, 36:1, 84-106, DOI:10.1080/02680939.2019.1686540

Szwecja

2009 reforma (wprowadzona 2012)–
obniżenie wieku uczniów podlegających
formalnej ocenie z 14 do 12/13 l. +
poszerzenie zakresu testów kompetencyjnych
w 12 r. ż. (ogólnokrajowych)

Powody: spadek wyników PISA, potrzeba
monitorowania uczniów/nauczycieli (szkoły
konkurują a oceny mają być wyznacznikiem
pracy nauczycieli), potrzeba klarownej
informacji zwrotnej dla rodziców,
„zaskoczenie” na późniejszych selektywnych
etapach

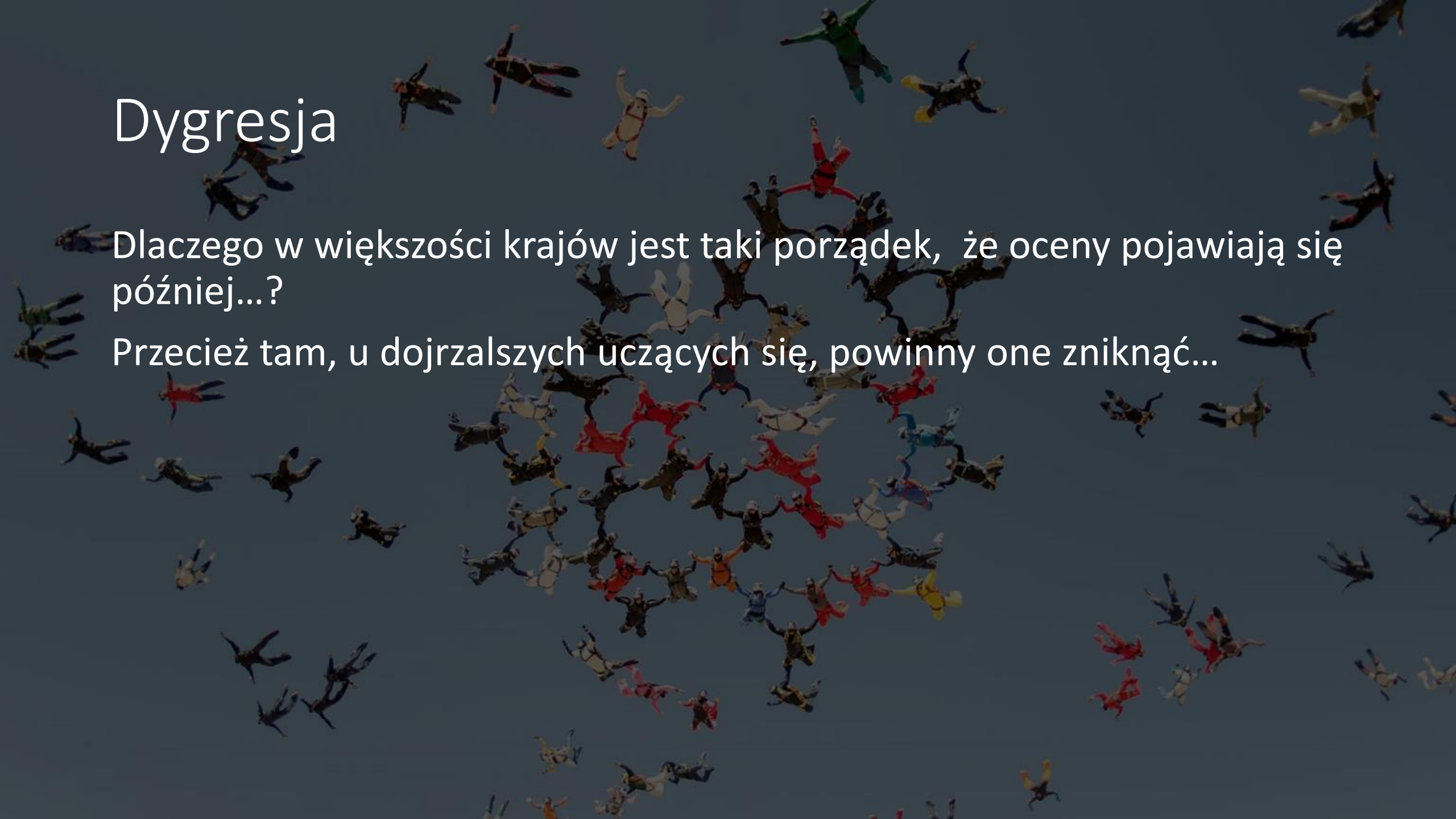
Stres - tylko w kontekście „przyzwyczajenia”
do ocen na późniejszym etapie



Dygresja

Dlaczego w większości krajów jest taki porządek, że oceny pojawiają się później...?

Przecież tam, u dojrzałych uczących się, powinny one zniknąć...



Model szwedzkiego badania



Bazujące na HBSC: Stres związany ze szkołą, samoocena osiągnięć w nauce, symptomy psychologiczne i psychosomatyczne, satysfakcja z życia

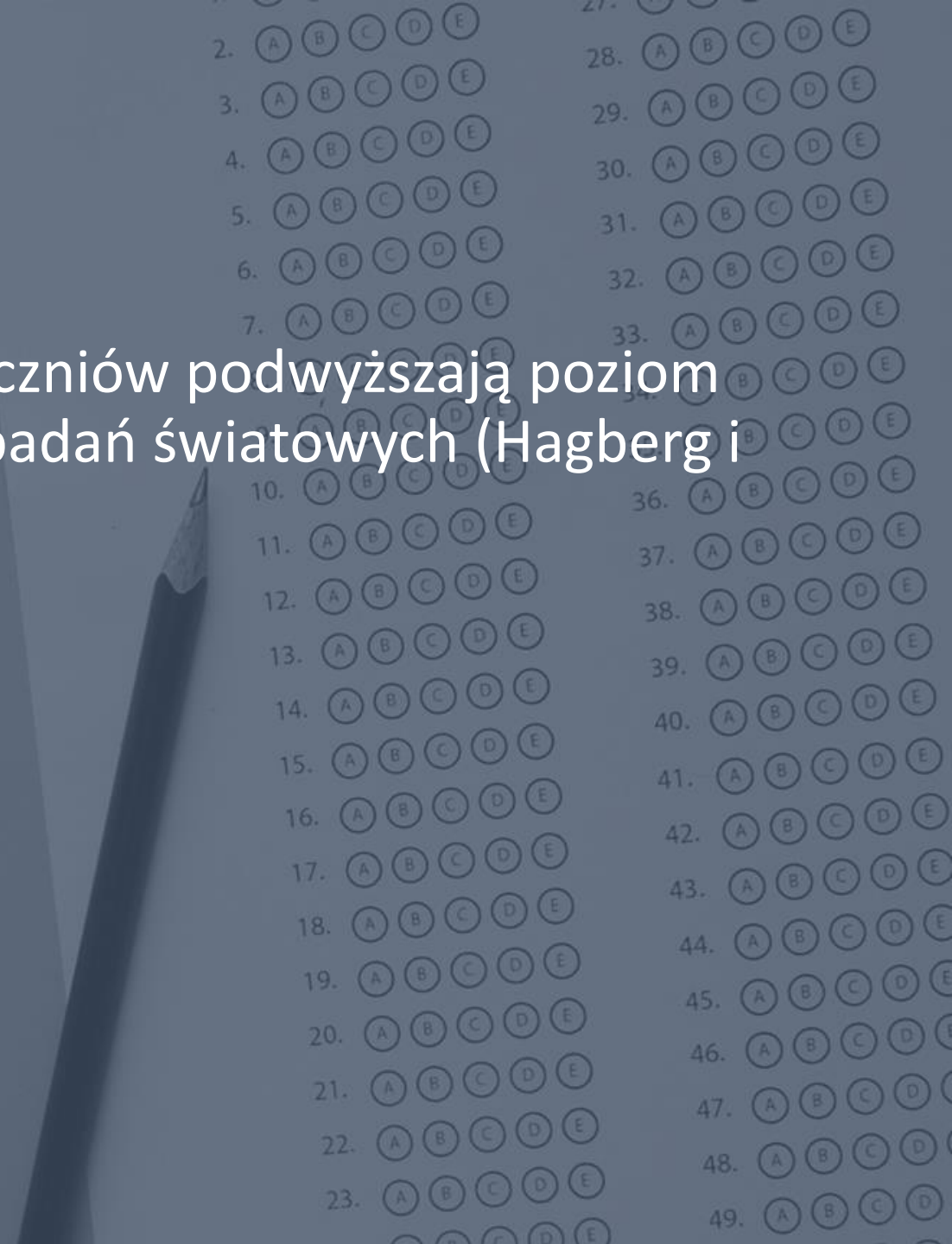
Reforma oceniania była przedmiotem dwóch badań jakościowych. Löfgren i Löfgren (2016) przeprowadzili wywiady z uczniami klas szóstych, które przyniosły zróżnicowane odpowiedzi. Niektórzy zgłaszali wyższy poziom motywacji, podczas gdy inni doświadczali większego stresu po otrzymaniu ocen. Wspólnym tematem było skłonność do porównywania ocen z kolegami, co mogło zwiększać presję wynikającą z osiągnięć w szkole. Na podstawie etnograficznego badania uczniów klas szóstych, Olovsson (2015) stwierdził, że wielu uczniów odczuwało większą presję wynikającą z ocen i narodowych testów, ale jednocześnie wykazywali większą dyscyplinę i motywację w nauce.

Co z tego wyszło?

Ustaliliśmy, że reforma zwiększyła stres i zmniejszyła samoocenę akademicką uczniów, a także dostarczyliśmy pewnych dowodów na to, że spowodowała ona również większą liczbę objawów psychosomatycznych oraz zmniejszoną satysfakcję z życia, co z kolei głównie wynikało z zwiększonego stresu i zmniejszonej samooceny. W przypadku satysfakcji z życia negatywny wpływ reformy był silniejszy u dziewcząt, zwiększając w ten sposób różnicę między płciami. Ogólnie rzecz biorąc, wyniki były słabsze w przypadku objawów psychosomatycznych jako wyniku, podczas gdy najbardziej solidne wnioski można wyciągnąć w odniesieniu do bezpośrednio związanych ze szkołą miar – stresu i samooceny – oraz satysfakcji z życia. Jedną z interpretacji tego zjawiska jest to, że zwiększone skupienie na ocenianiu ma głównie wpływ na wyniki bardziej bezpośrednio (tj. bezpośrednio związane ze szkołą) lub oceniające i "pozytywne" (tj. satysfakcję z życia), ale konsekwencje te nie są wystarczająco silne, aby bezpośrednio wpływać na objawy zdrowia somatycznego lub psychicznego.

Inny wymiar - testowanie

Generalnie testy „o wysoką stawkę” dla uczniów podwyższają poziom stresu i kosztów zdrowotnych – przegląd badań światowych (Hagberg i in., 2021)



Oceny jako podstawa porównań z innymi

Częsta informacja – sukcesy w nauce, to sukces w życiu

Kategoryzowanie młodych ludzi – miejsce w hierarchii społecznej oparte na zewnętrznej walidacji

Porównywanie się z innymi

Poczucie własnej wartości

Potencjalne konflikty – np. przy różnicowaniu kryteriów oceniania, dostosowywanie ich do możliwości

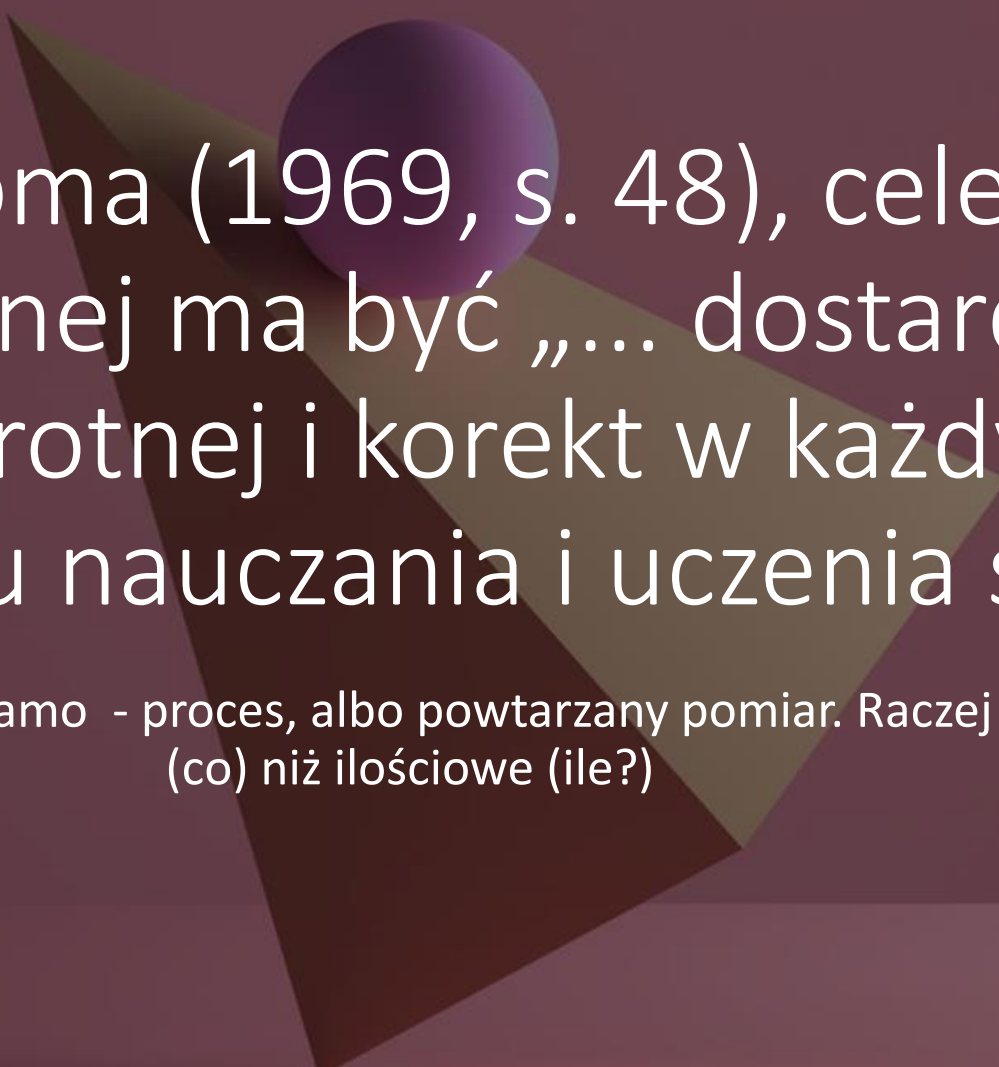
Oceny, w porównaniu do ocen nieformalnych lub formacyjnych, wprowadzają więcej możliwości dla uczniów, aby formalnie nie zdać, co może negatywnie wpływać na ich poczucie własnej wartości (Schraml et al. 2011; Giota and Gustafsson 2017; Ryan and Ryan 2005).

Dzieci ze środowisk defaworyzowanych i ich opowieść o ocenach

Kostet, I., Clycq, N., & Verschraegen, G. (2021). 'Maybe life can become easier because of my good grades': children's conflicting repertoires on aspirations and life chances. *Sociological Research Online*, 26(3), 581-600.

Wierzę w siebie. Jednak nie w taki sposób, że powiem, że stanę się lepszy od Messiego czy kogoś podobnego, ale po prostu... Ponieważ mam dość wysokie oceny, a dobrej firmy potrzebujesz dyplomu. Dlatego staram się uzyskać jak najlepsze oceny, jak to tylko możliwe, jak najlepsze... Jak najlepsze oceny, aby gdy będę chciał zatrudnić się w dobrej firmie, będę mógł pokazać moje kwalifikacje, a oni zobaczą "Aha, ma dobre oceny, zatrudnimy go". (12 lat, identyfikuje się jako Francuz, Marokańczyk i Belg)

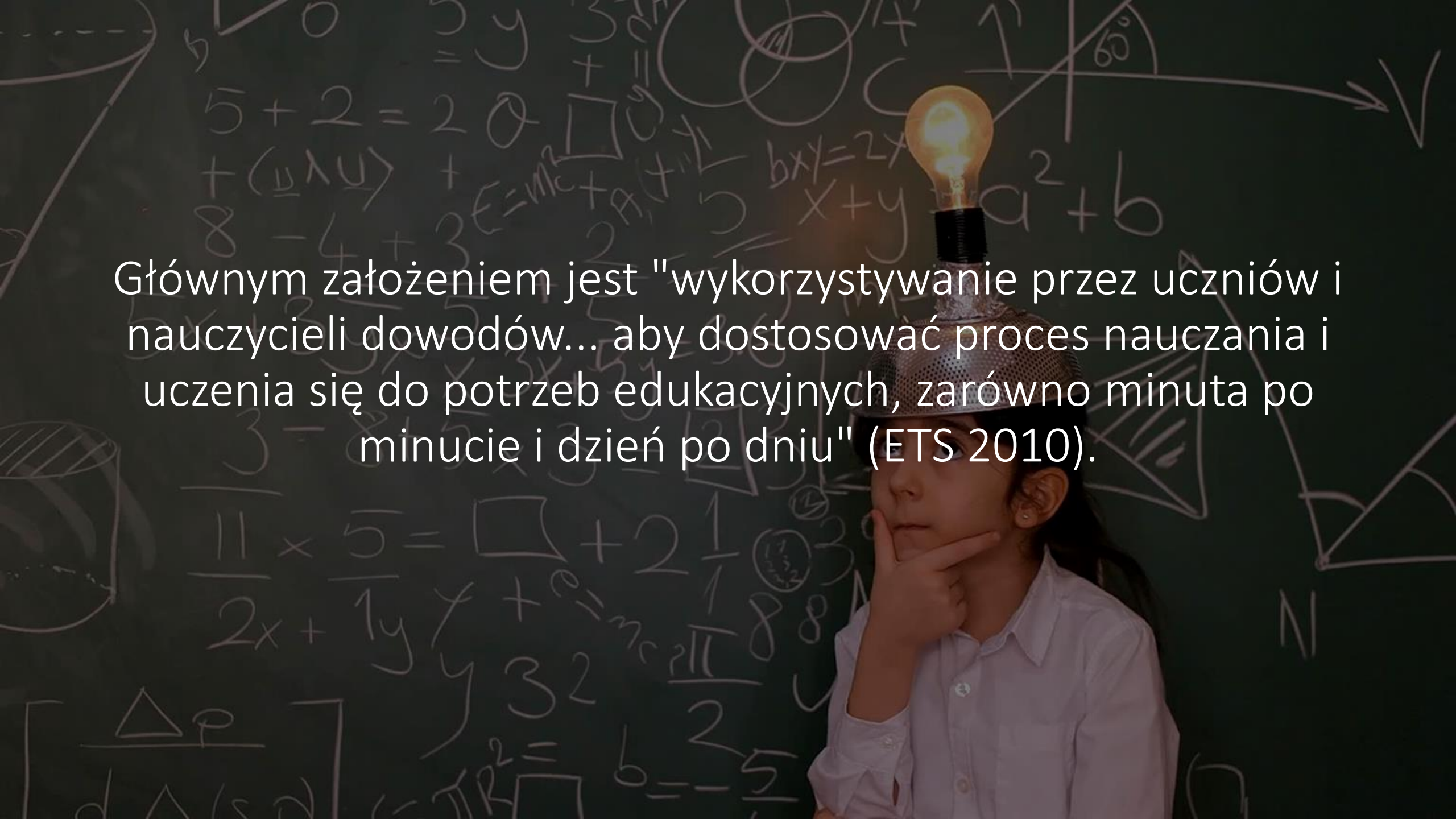




Według Blooma (1969, s. 48), celem oceny formatywnej ma być „... dostarczanie informacji zwrotnej i korekt w każdym etapie procesu nauczania i uczenia się”

Nie wszędzie rozumiane tak samo - proces, albo powtarzany pomiar. Raczej podejście jakościowe (co) niż ilościowe (ile?)

Assesement of learning czy
Assesement for learning?

A young girl with a glowing lightbulb on her head, symbolizing an idea or thought. She is wearing a white shirt and has her hand to her chin in a thinking pose. The background is a chalkboard filled with various mathematical equations and diagrams, including $5 + 2 = 2$, $E = mc^2$, $bxy = 2x$, $x + y$, $a^2 + b$, $11 \times 5 =$, $2x + 1y$, 32 , $6 = 2$, 5 , 60° , and N .

Głównym założeniem jest "wykorzystywanie przez uczniów i nauczycieli dowodów... aby dostosować proces nauczania i uczenia się do potrzeb edukacyjnych, zarówno minuta po minucie i dzień po dniu" (ETS 2010).

Często bezkrytycznie pozytywnie

- Ale:
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in education: principles, policy & practice*, 18(1), 5-25.
- Dunn, K. E., & Mulvenon, S. W. (2019). A critical review of research on formative assessments: The limited scientific evidence of the impact of formative assessments in education. *Practical assessment, research, and evaluation*, 14(1), 7.





Ocenianie kształtujące

Pięć kluczowych strategii obejmuje: dzielenie się oczekiwaniami dotyczącymi nauki (czyli wyjaśnianie i przekazywanie intencji nauki oraz kryteriów sukcesu), sprawdzanie (czyli prowadzenie efektywnych dyskusji klasowych, zadawanie pytań i zadań naukowych, które pozwalają uzyskać dowody nauki), udzielanie informacji zwrotnej, samoocenę (czyli aktywizację uczniów jako właścicieli własnego procesu nauki) oraz ocenę rówieśniczą (czyli aktywizację uczniów jako źródeł instruktażu dla siebie nawzajem).

Ile tych rzeczy musi być robione i w jaki sposób, żeby określić, że ktoś stosuje ocenianie kształtujące i zmierzyć jak ono działa? – do tego jeszcze wrócimy?

I jeszcze, jeżeli coś z tego działa, to który aspekt najbardziej?

Czy mamy dowody, że to lepiej działa?

- Black, P., and D. Wiliam. 1998. Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice* 5, no. 1: 7–74.
- *Być może byłoby pożądane... aby przegląd tego typu próbował dokonać metaanalizy badań ilościowych, które zostały opublikowane... Istnieją pojedyncze badania ilościowe, które biorą pod uwagę ocenę formatywną jako całość..., chociaż liczba badań o odpowiednim i porównywalnym wzorze metodologii ilościowej byłaby najwyżej rzędu 20. Jednakże, chociaż te [badania] są rygorystyczne w ramach swoich własnych struktur i celów,... podstawowe różnice między badaniami są takie, że jakiegokolwiek uogólnienia ich wyników miałyby niewielkie znaczenie (Black i William, 1998, s. 53).*
- Ciekawe są też wnioski dotyczące konkretnych badań z tych zestawień, np. że rodzaj informacji zwrotnej ma inny efekt w przypadku zadań różnego rodzaju, ale także cech uczniów (!!!). Wnioski? Czy istnieje uniwersalny system, dobry dla wszystkich i na każdym etapie?

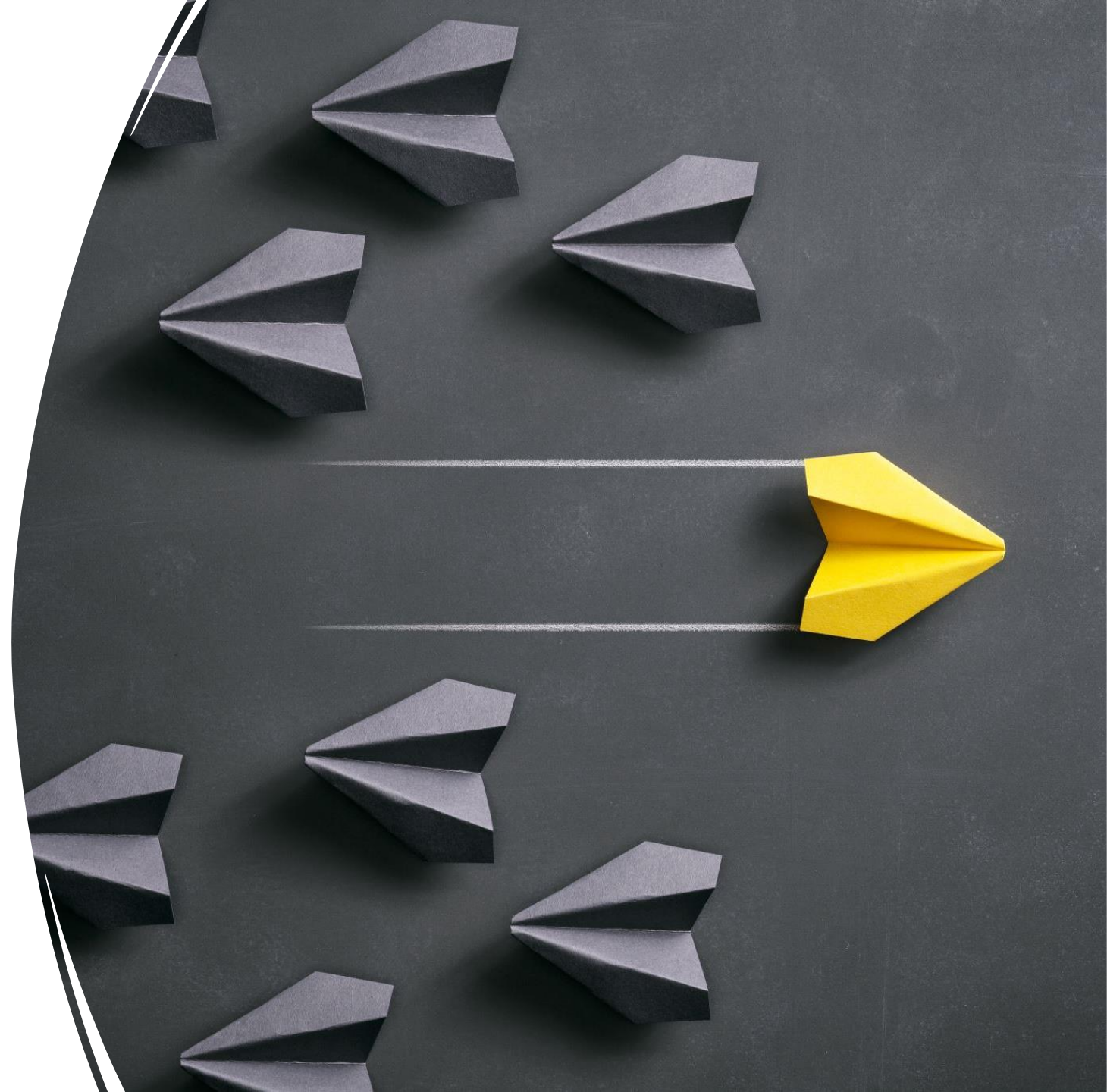
Do tego...

Każda strategia ma swoje techniki, a każdą technikę można zepsuć, nawet przez małe sprawy

Zatem:

Czy źle prowadzone ocenianie kształtujące może mieć gorsze efekty, niż „humanistyczne” obudowane dobrą relacją i komunikacją ocenianie z użyciem stopni?

Wydaje się, że tak...



„odwieczne obawy dotyczące
etycznej praktyki oceniania...
przyjmują nowy wymiar w
środowisku edukacji zdalnej”

(Abbott, Siskivic, Nogues i Williams
2000)

Ziarna do refleksji

1. Mechaniczne pozbycie się ocen nie rozwiąże wielu problemów, które ze sobą niosą (ale niektóre z nich odroczy)
2. Brak ocen przyniesie ze sobą potrzebę wytworzenia innych narzędzi, do spełnienia celów, którym służą oceny – pytanie czy szkoła bez czegoś, czy szkoła z czymś? (chyba, że jakieś cele znikną)
3. Etyczny i psychospołeczny aspekt oceniania każe zadawać wciąż pytania „Jak to robić” na każdym szczeblu (od organizacji systemu egzaminów po poziom konkretnej klasy szkolnej i konkretnego nauczyciela)
4. Reformy w systemie oceniania muszą być ekologiczne - ingerują one w postawy i doświadczenia wielu ludzi – np. rodziców

Ziarna do refleksji

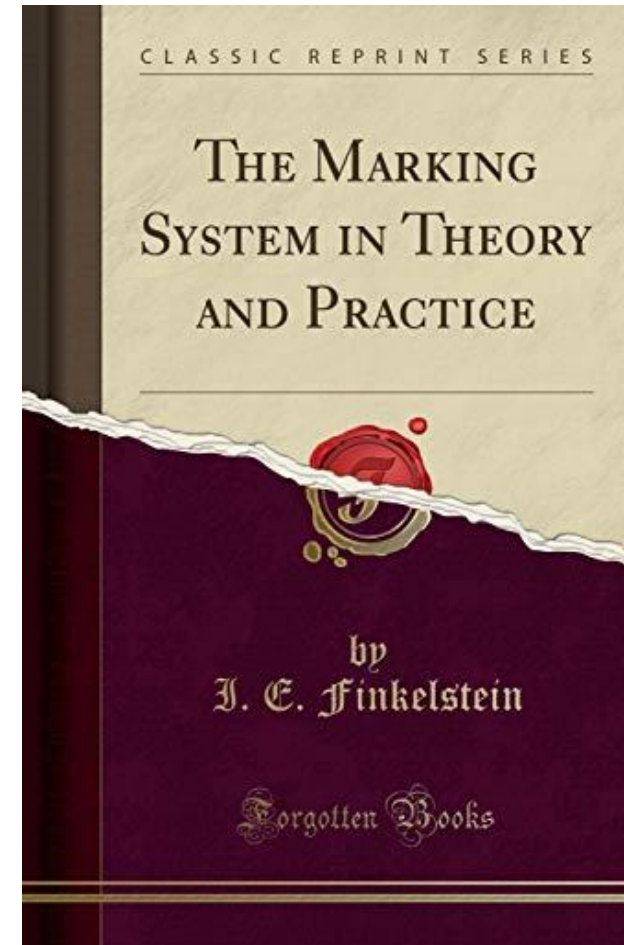
5. Bardzo istotne jest osadzenie systemu oceniania i testowania w kontekście – co z niego wynika?
6. Istotny jest poziom mikro – czy konkretny nauczyciel przygląda się temu jak ocenia, czy jest sprawiedliwy, jak komunikuje oceny (także „złe”) – element kształcenia i wspierania nauczycieli
7. Alternatywne systemy oceniania np. ocenę opisową, ocenianie kształtujące, też można zepsuć, tak, że przyniosą wiele szkód. Dowody na to, że jakiś sposób oceniania jest lepszy w każdym kontekście nie istnieją.
8. Może ważniejsze nie jest uniwersalne rozwiązanie a elastyczność?

I jeszcze do tego, udajemy, że
mamy dowody na to, że to
czego chcemy jest lepsze...

To nie nowe rozważania...

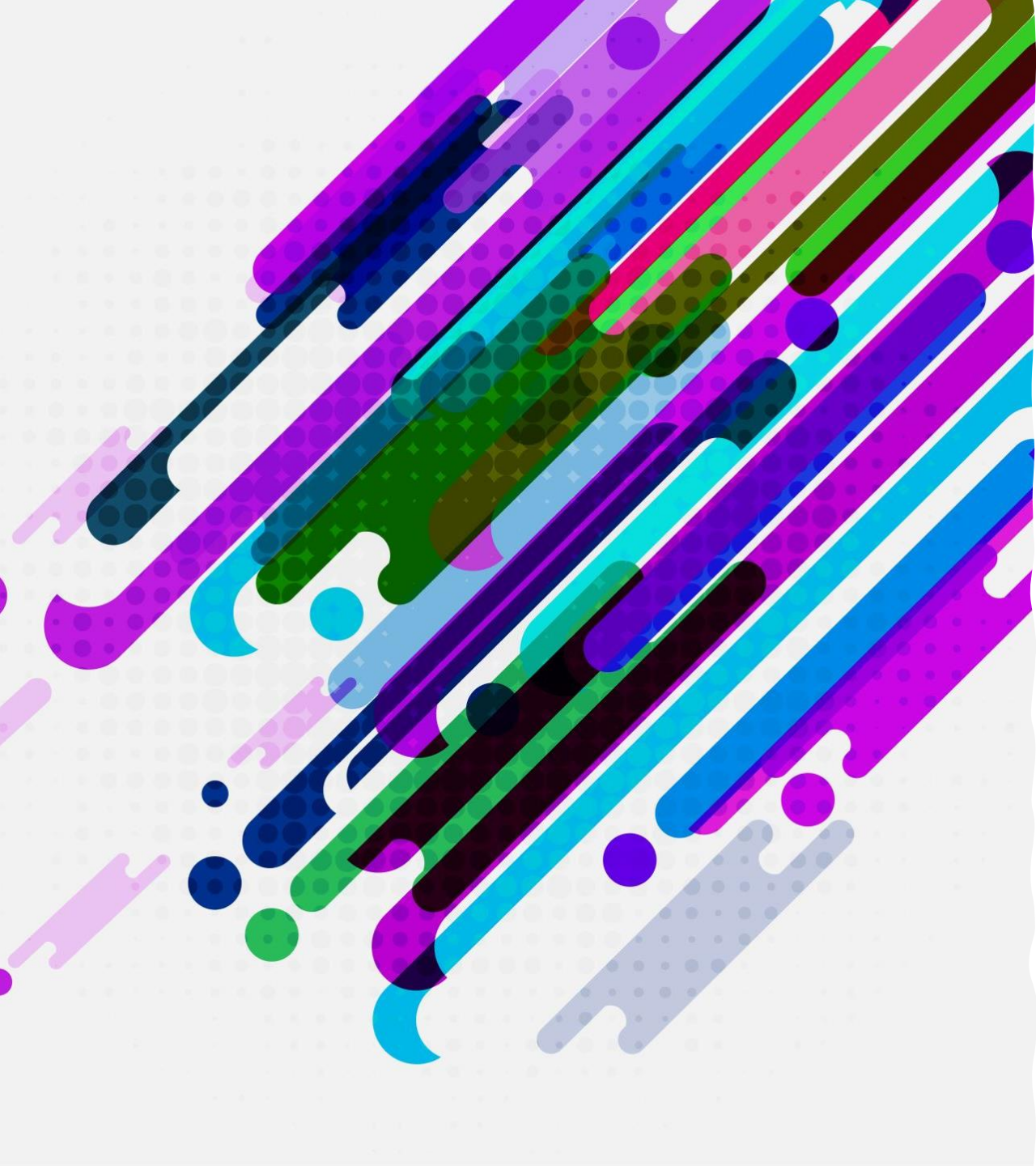
Kiedy rozważamy praktycznie uniwersalne stosowanie w wszystkich placówkach edukacyjnych systemu ocen, czy to w postaci liczb czy liter, aby wskazać osiągnięcia szkolne uczniów w tych placówkach, i kiedy pamiętamy, jak ogromny nacisk kładziony jest zarówno przez nauczycieli, jak i uczniów na te oceny jako rzeczywiste miary osiągnięć, nie możemy się nie dziwić ślepeму wierzeniu w niezawodność systemów ocen.

—I. E. Finkelstein (1913)



Wciąż te same dobre pytania, bez pewnych odpowiedzi

W ciągu ostatniej dekady poważna uwaga została skupiona na pytaniach takich jak: Co naprawdę powinna oznaczać ocena? Czy powinna być oparta na zdolnościach, osiągnięciach, a może nawet na zapale i entuzjazmie? Jakie są najlepsze symbole do reprezentowania zdolności czy osiągnięć? Jak rozkładają się oceny przyznawane przez różnych nauczycieli czy różne szkoły? Czy poprzez prezentację rozkładów ocen lub formalne nauczanie teorii oceniania można zwiększyć sprawiedliwość i wiarygodność ocen? Czy uczniowie mają tendencję do osiągania podobnej pozycji pracując z różnymi nauczycielami w tej samej szkole, czy też utrzymują swoje względne pozycje przechodząc z klasy do klasy lub ze szkoły do college'u? (Finkelstein, 1913)



Te pytania dedykuję
wszystkim, którzy zawsze
w edukacji mają gotowe
i pewne odpowiedzi... (a
to szybko powiększająca
się grupa)

Paradoksalnie, czym się więcej o
czymś wie w edukacji, tym mniej
jesteśmy pewni rozwiązań