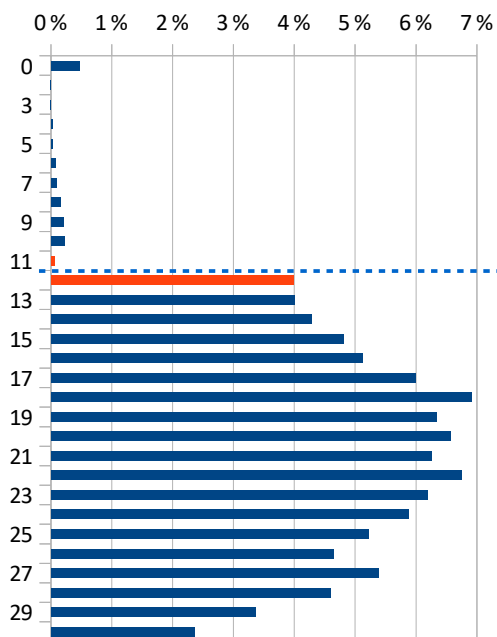


Kolik žáků zachránili hodnotitelé na školách před propadnutím ve slohu?

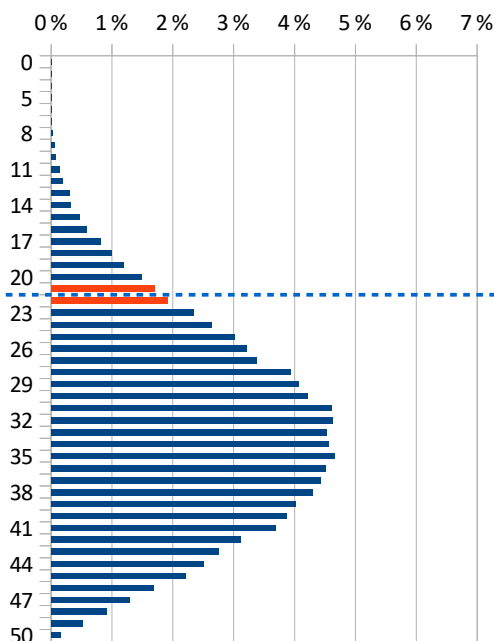
Rozložení úspěšnosti ve slohu



2016



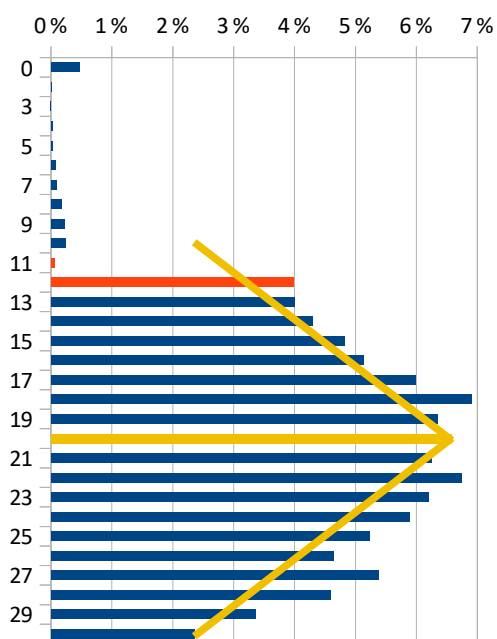
Rozložení úspěšnosti v testu



Oba diagramy znázorňují, kolik procent žáků (0 % až 7 % – viz vodorovná osa) získalo ve slohu, případně v testu počet bodů uvedený na svislé ose (0–30 bodů ve slohu, 0–50 bodů v testu). Červeně jsou znázorněny sloupce odpovídající hraničnímu výsledku: například 11 bodů ve slohu znamenalo „neprospěl“, 12 bodů znamenalo „prospěl“. V testu šlo o hodnoty 21, resp. 22 bodů.

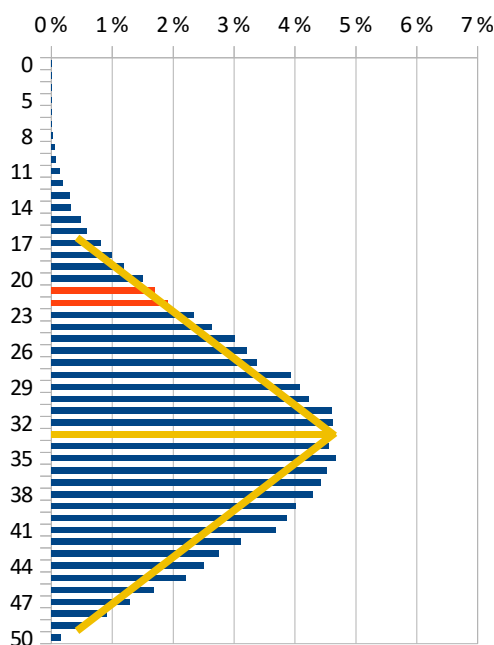
Tytéž diagramy jsme potom doplnili STEJNOU dvojicí okrových čar (>) symetrických podle vodorovné osy (—). Procházela mediánem. Ve slohu to znamená, že méně, resp. více než 20 bodů získaly stejné počty žáků. V testu to znamená, že méně, resp. více než 33 bodů získaly stejné počty žáků. V diagramu je tedy součet délek sloupců nad osou šipky stejný jako součet délek sloupců pod ní.

Rozložení úspěšnosti ve slohu



2016

Rozložení úspěšnosti v testu



Podle důvodové zprávy k novele školského zákona, jež vrátila hodnocení slohů ze škol do centra, *došlo v letech 2013 a 2014 k nárůstu počtu písemných prací maturujících žáků, které byly učiteli z kmenových škol hodnoceny takovým nejnižším možným bodovým hodnocením, aby žák ještě u této dílčí zkoušky uspěl a nemusel zkoušku opakovat.*

Míru tohoto přilepšení žákům pod hranicí propadnutí vyjadřuje „schod“ mezi červenými čarami odpovídajícími výsledkům 11 bodů (neprospěl) a 12 bodů (prospěl). Z diagramu je ovšem také patrné, že přilepšení se mohlo týkat přibližně 4 % maturantů, což je vzhledem k jeho podílu na celku zanedbatelné. Přilepšení je však stěží prokazatelné vzhledem k velké neurčitosti hodnotících kritérií¹. Například „přesahy“ okrové šipky vykazují u slohu i u testu podobnou nesymetrii, v zásadě velmi mírnou. V testu však pro vliv subjektivních zásahů hodnotitelů nebyl žádný prostor.

Podle zpracovaných dat za jaro 2016 nebyl k přesunu hodnocení ze škol do centra žádný zásadní důvod.

Zmíněná důvodová zpráva k novele školského zákona uvádí, že na kontrolním vzorku z roku 2014 vzrostla po přehodnocení neúspěšnost z hodnoty 1,8 % na 10,1 %. **To byl také hlavní argument pro návrat hodnocení slohů ze škol do centra. Důvodová zpráva ovšem neuvádí ani postup výběru hodnotitelů, kteří prováděli přehodnocení, ani instrukce, které hodnotitelé dostali. Nepopisuje dokonce ani velikost vzorku a jeho další vlastnosti.** Předkladatel zákona tento argument přesto použil a naši zákonodárci ho přijali.

Diagramy na předcházející straně zachycují přibližně 62 tisíc žáků, kteří psali sloh i test z češtiny. Na jaře 2016 z nich ve slohu neuspělo 1,35 %. Kdyby měl argument z důvodové zprávy platit i v roce 2016 pro celý soubor, museli by hodnotitelé na školách nadržováním zachránit před propadnutím 8,75 % všech žáků. **Kde tito žáci, kteří prý správně měli ve slohu propadnout, v diagramu jsou?** Znamená to snad, že je máme hledat i mezi žáky s průměrným, či dokonce nadprůměrným výsledkem? Nebo snad byly výsledky žáků ve slohu nadhodnoceny plošně?

Jediný konkrétní argument pro přesun hodnocení ze škol do centra uvedený v důvodové zprávě má pro celek nulovou vypovídací hodnotu a je vlastně – jako takový – spíše k smíchu.

Korektní experiment s přehodnocením by měl dodržet mimo jiné následující pravidla:

Hodnotitelé provádějící přehodnocení **nesmějí vědět, co je cílem experimentu.** Nesmějí tudíž obdržet k přehodnocení žádné jiné instrukce než ty, které mají k dispozici při hodnocení běžném. Cíl experimentu nesmějí odhalit ani z převládající úrovně hodnocených písemných prací. Proto musejí dostat k přehodnocení jak práce slabé, tak práce průměrné i nadprůměrné. **Práce musejí být vybrány náhodně a vzorek má být dostatečně velký,** aby bylo možné zobecnit získaný výsledek na celý soubor. Zhruba šestinásobný nárůst neúspěšnosti se totiž může projevit i na vzorku pouhých deseti prací. **Náhodný musí být i samotný výběr hodnotitelů.** Mezi hodnotiteli na školách byli i lidé, kteří se veřejně zasazovali o návrat centrálního hodnocení. Pokud by převládli v týmu, který prováděl přehodnocení, **už jenom kvůli tomuto nedostatku by nešlo objektivitu celého experimentu úspěšně hájit.**

¹ Například v kritériu 3A týkajícím se výstavby větných celků a nemotivovaných odchylek od pravidelné větné stavby (jejichž výskyt je charakterizován slovy *ve vysoké míře / ve větší míře / občas / místy / jen ojedinele / nikdy*) má být hodnotiteli posuzován též vliv těchto nedostatků na čtenářský komfort adresáta: *je zásadní / existuje / je občasný / existuje místy / neexistuje / komfort není narušován.*