Scientia in educatione, 14(1), 2023
http://www.scied.cz
ISSN 1804-7106

# Obsah

## Výzkumné studie





Nakladatelství Karolinum

# Úroveň bádateľských zručností žiakov 6. ročníka základných škôl

## Level of inquiry skills among 6th grade primary school pupils

Elena Čipková[1], Michael Fuchs[1], Dominik Šmida[1,*]

[1] Katedra didaktiky prírodných vied, psychológie a pedagogiky, Prírodovedecká fakulta Univerzita Komenského v Bratislave, Ilkovičova 6, 84215 Bratislava, Slovensko; smida8@uniba.sk

Primárnym cieľom prírodovedného vzdelávania je rozvíjať prírodovednú gramotnosť žiaka tak, aby disponoval širokým spektrom vedomostí, postojov, kompetencií a zručností, ktoré mu umožňujú využívať a uplatňovať poznatky a postupy spojené s vedeckým skúmaním v každodennom živote pri riešení rôznych situácií. Významnou zložkou prírodovednej gramotnosti sú bádateľské zručnosti, prostredníctvom ktorých sa pri riešení problémov môžu uplatňovať metódy a postupy kopírujúce povahu práce vedcov. Žiakom zároveň umožňujú získavať nové poznatky a porozumieť prírodným konceptom, na základe ktorých môžu hlbšie preniknúť do spôsobu fungovania okolitého sveta. V predloženom príspevku predstavujeme výsledky realizovaného výskumu zameraného na hodnotenie úrovne bádateľských zručností žiakov šiesteho ročníka nižšieho sekundárneho vzdelávania (ISCED 2). Získané údaje naznačujú, že žiaci participujúci na výskume disponujú pomerne nízkou úrovňou bádateľských zručností, čo môže výrazne ovplyvniť ich schopnosť vyhľadávať a objavovať nové poznatky z oblasti prírodných vied a v konečnom dôsledku mať negatívny dopad na ich celkovú úroveň prírodovednej gramotnosti.

Klíčová slova:
bádateľské zručnosti, nižšie sekundárne vzdelávanie, prírodovedná gramotnosť.

Zasláno 10/2022
Revidováno 3/2023
Přijato 6/2023

The primary aim of science education is to develop the pupils' scientific literacy that they have a wide range of knowledge, attitudes, competencies, and skills that enable them to use and apply knowledge and procedures associated with scientific inquiry when solving various situations in daily life. An important component of scientific literacy is inquiry skills, through which methods and procedures copying the work of scientists can be applied in solving problems. At the same time, they allow pupils to gain new knowledge and understand natural concepts, based on which they can input more deeply into the way the world works. In the submitted contribution, we present the results of the conducted research aimed at evaluating the level of inquiry skills among sixth-grade pupils of lower secondary education (ISCED 2). The obtained data indicate that the pupils who participated in the research have a relatively low level of inquiry skills, which can significantly affect their ability to search and discover new knowledge in the field of natural sciences and ultimately have a negative impact on their overall level of scientific literacy.

Key words:
inquiry skills, lower secondary education, scientific literacy.

Received 10/2022
Revised 3/2023
Accepted 6/2023

# 1 Úvod

Vo všeobecnosti sa za základný cieľ prírodovedného vzdelávania považuje rozvoj prírodovednej gramotnosti (DeBoer, 2000; Liu, 2013). V kontexte prírodovedného vzdelávania môžeme prírodovednú gramotnosť vymedziť ako rozvíjanie kompetencií spojených s využívaním vedeckých poznatkov a zručností založených na získavaní dôkazov, ktoré majú význam pre každodenný život pri riešení osobne náročných, ale z hľadiska vedy zmysluplných problémov (Holbrook & Rannikmae, 2009). Prírodovedná gramotnosť tak predstavuje požiadavku, že určité zvládnutie vedy je základným predpokladom nie len pre vzdelávanie, ale aj pre každodenný život v spoločnosti. Harlen (2001) vymedzila tri zložky, ktoré sa podieľajú na utváraní prírodovednej gramotnosti. Ide o prírodovedné predstavy, prejavy vedeckého postoja k realite a spôsobilosti vedeckej práce. Najobsiahlejšiu zložku prírodovednej gramotnosti predstavujú práve spôsobilosti vedeckej práce, ktoré sú niektorými autormi (napr. Balogová & Ješková, 2016; O'Connor & Rosicka, 2020; Song, 2016) stotožňované s bádateľskými zručnosťami. Rozvoj tohto súboru zručností predstavuje kľúčový prvok aj pre formovanie zvyšných dvoch zložiek prírodovednej gramotnosti (Lou et al., 2015; Rezba et al., 2003). Podľa Wenninga (2007) práve osvojenie si procesov vedeckého bádania umožňuje naplniť cieľ prírodovedného vzdelávania, ktorým je rozvoj prírodovednej gramotnosti. Publikované štúdie (napr. Bellová et al., 2018; Kotuľáková, 2020; Miškovičová et al., 2009; Mukti et al., 2019) dlhodobo poukazujú na pomerne nízku úroveň prírodovednej gramotnosti žiakov, ako na Slovensku, tak aj v iných krajinách. Na klesajúci trend v úrovni prírodovednej gramotnosti slovenských žiakov poukazujú aj výsledky meraní PISA (Miklovičová & Valovič, 2019) a TIMSS (TIMSS, 2019), v ktorých žiaci, na rozdiel od žiakov v ČR, dosiahli výsledky pod priemerom krajín OECD (Mullis et al., 2020; Schleicher, 2019). Podľa Firmana (2007) nízka úroveň prírodovednej gramotnosti priamo súvisí s charakterom


https://doi.org/10.14712/18047106.2276
Nakladatelství Karolinum

vzdelávania, ktoré v nedostatočnej miere kladie do popredia aktivity, ktoré by viedli k zvýšeniu úrovne bádateľských zručností. Úroveň bádateľských zručností totiž výraznou mierou determinuje schopnosť žiakov osvojiť si obsah prírodovedného vzdelávania vychádzajúceho z vedeckého poznania (Harlen, 2014). Inovovaný Štátny vzdelávací program na Slovensku (ŠPÚ, 2015) pritom presadzuje, aby žiaci v rámci vyučovania prírodných vied dostali príležitosť bádať, objavovať, rozvíjať si bádateľské zručnosti a formovať si celkové poznanie v oblasti prírodných vied. S ohľadom na túto skutočnosť je potrebné vo výchovno-vzdelávacom procese dôsledne venovať pozornosť skúmaniu a hodnoteniu úrovne bádateľských zručností žiakov a hľadaniu vhodných stratégií na ich cielený rozvoj.

## 2 Bádateľské zručnosti

Prírodovedné vzdelávanie kladie dôraz nie len na nadobúdanie a porozumenie prírodovedným konceptom (DeBoer, 2000), ale aj na osvojenie si spôsobov vedeckej práce, prostredníctvom ktorých môžu žiaci získať nové vedomosti a porozumieť skutočnej povahe vedy (Wang et al., 2015). Do popredia sa tak dostalo vzdelávanie založené na bádaní. V odbornej literatúre sa stretávame s rôznymi vymedzeniami pojmu bádanie. Minárechová (2014) a Colburn (2000) definujú bádanie ako otvorené praktické aktivity smerované na žiaka, prostredníctvom ktorých môžeme podľa Llewellyna (2013) rozvíjať kritické myslenie žiakov, ich vedomosti, zručnosti, postoje a návyky, ktoré je možné aplikovať nie len v školskom prostredí, ale aj v bežnom živote. Kireš et al. (2016) charakterizujú bádanie ako aktívne skúmanie žiakov s cieľom nachádzania odpovedí na otázky, ktoré ich zaujímajú. Harlen (2013) dokonca bádanie prirovnáva k investigatívnemu vyšetrovaniu, ktoré umožňuje žiakom prostredníctvom priamej interakcie s okolím a získavaním dôkazov spoznať a pochopiť okolitý svet. Pri bádaní sa tak aktivita presúva na žiaka, pričom učiteľ vystupuje ako facilitátor a usmerňuje prácu žiakov (Čepičková, 2013). Pre realizáciu uvedeného spôsobu vyučovania prírodných vied je nevyhnutné, aby žiaci disponovali určitou úrovňou bádateľských zručností, ktoré im umožnia postupovať pri bádaní obdobne ako vedcom.

Bádateľské zručnosti predstavujú základný pilier slúžiaci žiakom na pochopenie sveta vedy a prírody (O'Connor & Rosicka, 2020). Millar a Driver (1987) definujú bádateľské zručnosti v kontexte vzdelávania ako súbor zručností, ktoré sú využívané pri realizácii bádania v procese vyučovania prírodných vied, pričom Kireš et al. (2016) považujú bádateľské zručnosti za súbor tzv. *„soft-skills“* a špecifických zručností, ktoré sú charakteristické pre prácu vedca a nevyhnutné pre realizáciu akejkoľvek vedeckej činnosti spojenej so skúmaním (Stone, 2014). Prostredníctvom týchto zručností si žiaci rozvíjajú schopnosť klásť otázky a hľadať vhodné odpovede (NRC, 1996), pričom umožňujú žiakom zapojiť sa do procesu vzdelávania v súlade s princípmi konštruktivizmu (ŠPÚ, 2015).

V literatúre sa môžeme stretnúť s rôznymi podobami klasifikácií bádateľských zručností. Podľa Ješkovej et al. (2016a) tieto klasifikácie môžeme rozdeliť do dvoch základných rámcov. Prvý rámec tvoria klasifikácie, ktoré sú odvodené od jednotlivých stupňov bádania. Do tohto rámca spadá napríklad klasifikácia podľa Fradda et al. (2001), ktorí rozdelili bádateľské zručnosti na základe fáz bádateľského cyklu na šesť skupín, a to formulovanie problému, plánovanie, implementáciu, vyvodzovanie záverov, zdieľanie výsledkov a aplikovanie. Druhý rámec predstavujú klasifikácie navrhnuté s ohľadom na vekovú kategóriu žiakov. Príkladom je klasifikácia od Wenninga (2005), ktorý rozdelil jednotlivé bádateľské zručnosti podľa veku a intelektuálnej úrovne žiakov. Vymedzil tak štyri skupiny zručností, a to elementárne, základné, integrované a pokročilé. Uvedená klasifikácia bola autorom v rámci ďalšieho výskumu doplnená o stredne pokročilé a kulminujúce bádateľské zručnosti, pričom vymedzenie jednotlivých kategórií zručností nie je rigidné a je možné ich prispôsobiť aktuálnym možnostiam žiakov (Wenning, 2010). Prehľad bádateľských zručností uvádzame v tab. 1.

Podľa Özgelena (2012) bádateľské zručnosti priamo súvisia aj s kognitívnym rozvojom osobnosti žiaka v oblasti prírodných vied. Tieto zručnosti poskytujú žiakom podporu pri myslení, uvažovaní, hodnotení a riešení konkrétnych situácií. Výskum Suryawati a Osmana (2017) naznačuje, že čím vyššou úrovňou bádateľských zručností žiaci disponujú, tým vyššia je aj úroveň ich kognitívneho rozvoja v oblasti prírodných vied. Úroveň bádateľských zručností žiakov koreluje s realizáciou bádateľských aktivít v prírodovednom vzdelávaní (Ergül et al., 2011). Realizované výskumy však poukazujú na výrazné rozdiely v úrovni zručností žiakov medzi jednotlivými krajinami (napr. Burns et al., 1985; Šmida & Čipková, 2021; Tan, 1996) a opisujú rôzne problematické aspekty ich rozvoja vo vzdelávaní (napr. Deters, 2005; Cheung, 2007). Uvedenú skutočnosť zdôvodňujú tým, že častokrát učitelia síce bádanie integrujú do vzdelávacích plánov v súvislosti s platnými kurikulárnymi dokumentmi, reálne ho však vo svojej výučbe nevyužívajú. Výskumy úrovne bádateľských zručností žiakov preto môžeme vnímať aj v kontexte toho, či dochádza k napĺňaniu princípov prírodovedného vzdelávania založeného na konštruktivizme.

V súčasnosti sú dostupné údaje z medzinárodných porovnávacích testovaní PISA alebo TIMSS, ktoré sa zameriavajú na posúdenie úrovne prírodovednej gramotnosti (Mullis et al., 2020; Schleicher, 2019).

**Tab. 1:** Klasifikácia bádateľských zručností (Wenning, 2010 – upravené)

| **Elementárne zručnosti** | **Základné zručnosti** | **Stredne pokročilé zručnosti** | **Integrované zručnosti** | **Kulminujúce zručnosti** | **Pokročilé zručnosti** |
|---|---|---|---|---|---|
| – pozorovať<br>– formulovať koncepty<br>– odhadovať<br>– formulovať závery<br>– komunikovať výsledky<br>– klasifikovať výsledky | – predpovedať<br>– vysvetliť vzťahy<br>– odhadovať<br>– získavať a spracovávať údaje<br>– formulovať vedecké a logické vysvetlenia na základe dôkazov<br>– rozpoznať a analyzovať alternatívne vysvetlenia a modely | – merať<br>– zbierať a zaznamenávať údaje<br>– zostrojiť tabuľku a graf<br>– naplánovať a realizovať vedecké bádanie<br>– identifikovať a kontrolovať premenné<br>– využívať technológiu a matematiku v procese bádania<br>– vysvetliť vzťahy | – metricky merať<br>– vytvárať empirické zákony na základe logických dôkazov<br>– naplánovať a riadiť vedecké bádanie<br>– využívať technológiu a matematiku v procese bádania | – zhromaždiť, hodnotiť a interpretovať dáta<br>– konštruovať argumenty založené na dôkazoch<br>– hodnotiť na základe dôkazov<br>– objasniť hodnoty vo vzťahu k prírodným zákonom a občianskym právam<br>– vzájomne spolupracovať | – tvoriť hypotetické vysvetlenia<br>– analyzovať a hodnotiť vedecké argumenty<br>– tvoriť predpovede prostredníctvom deduktívneho myslenia<br>– revidovať hypotézy a predpovede na základe nových dôkazov<br>– riešiť zložité problémy z bežného života |

sofistikovanosť intelektuálnych procesov

nižšia → vyššia

Výsledky týchto testov ale nepoukazujú na úroveň bádateľských zručností. Testovanie TIMSS sa zameriava iba na posúdenie úrovne vedomostí žiakov štvrtého ročníka základnej školy a z výsledkov PISA testovania nie je možné jednoznačne určiť úroveň jednotlivých bádateľských zručností žiakov. Z toho dôvodu je potrebné v rámci výskumu zamerať pozornosť aj na samotné bádateľské zručnosti žiakov. Na ich meranie sa najčastejšie využívajú testy (napr. Kruit et al., 2018; Shahali & Halim, 2010; Temiz, 2020; Tosun, 2019; Wenning, 2006, 2007 a pod.), ktoré umožňujú získať objektívne výsledky o výkone žiakov pri riešení úloh vyžadujúcich si bádateľsky prístup. Podrobnejšie údaje o spôsoboch, ktorými žiaci dokážu využívať bádateľské zručnosti v praxi, môžeme získať aj ich pozorovaním pri realizácii bádateľských aktivít (napr. Hairida, 2016; Mulyeni et al., 2019) alebo prostredníctvom analýzy rôznych nástrojov formatívneho hodnotenia (Harlen, 2013), ktoré nám umožňujú presnejšie identifikovať chyby žiakov v procese učenia sa a osvojovania si zručností (Opara & Oguzor, 2011). V snahe detailnejšie analyzovať jednotlivé zručnosti sa v súčasnej dobe kladie dôraz aj na analýzu jednej či dvoch konkrétnych zručností, ktoré sú posudzované na základe rôznych kritérií (napr. Kireš & Jurková, 2021; Nejedlý & Vojíř, 2022). Mnohé výskumy zároveň poukazujú na rozdiely medzi úrovňou osvojených bádateľských zručností a pohlavím žiakov (napr. Guevara, 2015; Ješková et al., 2021; Zeidan & Jayosi, 2015), avšak pri iných výskumoch tento rozdiel nebol signifikantne preukázaný (napr. Ješková et al., 2016a; Šmida & Čipková, 2021).

## 3 Cieľ výskumu a výskumné otázky

Pre rozvoj bádateľských zručností je nevyhnutné vo vzdelávaní využívať také metódy a postupy, ktoré žiakom umožnia kopírovať prácu vedcov (Windschitl, 2000), povedú k prehĺbeniu žiackych vedomostí z oblasti prírodných vied (Harrison, 2014) a umožnia u žiakov formovať si pozitívne postoje k prírode a k samotnej vede (Topalsan, 2020). Aby sme však dokázali cielene implementovať bádateľské aktivity do vyučovania, je potrebné poznať, akou úrovňou zručností žiaci aktuálne disponujú. Podľa Weninga

(2007) hodnotenie aktuálnej úrovne bádateľských zručností predstavuje základ pre posudzovanie úspešnosti vzdelávania a tvorbu kurikulárnych dokumentov. V súčasnosti sa na Slovensku pripravuje kurikulárna reforma základnej školy, v rámci ktorej bude obsah vzdelávania na základnej škole usporiadaný do troch cyklov, pričom tretí cyklus bude pokrývať šiesty až deviaty ročník. Navrhované zmeny v prírodovednej oblasti presúvajú ťažisko vzdelávania z odovzdávania vedomostí na rozvoj prírodovedného poznania žiaka prostredníctvom aplikácie metód a postupov objektívneho a systematického skúmania. Z toho dôvodu cieľom predloženého výskumu bolo zistiť úroveň vybraných bádateľských zručností žiakov šiesteho ročníka základných škôl na Slovensku. V súlade s cieľom výskumu sme si stanovili nasledovné výskumné otázky:

- Aká je úroveň vybraných bádateľských zručností žiakov šiesteho ročníka základných škôl na Slovensku?
- Aká je úroveň vybraných bádateľských zručností žiakov šiesteho ročníka základných škôl v závislosti od pohlavia?

# 4 Metodológia

Výskum sa zameriaval na zistenie úrovne bádateľských zručností žiakov šiesteho ročníka základných škôl. S požiadavkou na účasť vo výskume sme oslovili všetky plnoorganizované základné školy na Slovensku prostredníctvom e-mailu, pričom početnosť výskumného súboru ovplyvňovala ochota učiteľa sprístupniť administrovaný test žiakom v priebehu vyučovacej hodiny. Pre získanie potrebných údajov sme využili výskumný nástroj vlastnej konštrukcie. Výskumný nástroj bol školám administrovaný elektronicky prostredníctvom platformy Google Forms v období od januára do marca 2022. Účasť škôl na výskume bola dobrovoľná.

## 4.1 Výskumný súbor

Vzhľadom na stanovený výskumný cieľ sme využili dostupný výber výskumného súboru. Výskumu sa celkovo zúčastnilo 891 žiakov šiesteho ročníka základných škôl situovaných v rôznych regiónoch Slovenska. Výskumný súbor pozostával zo 423 chlapcov (47,5 %) a 468 dievčat (52,5 %), Priemerná známka žiakov na polročnom hodnotení z biológie bola 2,4. Vzdelávanie žiakov prebiehalo v súlade s inovovaným Štátnym vzdelávacím programom (ŠPÚ, 2015), ktorý v rámci vyučovania prírodných vied kladie do popredia bádanie, ako jeden zo spôsobov získavania nových vedomostí a zručností. Je ale potrebné si uvedomiť, že vzdelávanie na Slovensku v rámci základných škôl môže prebiehať v rôznych podmienkach a môže sa líšiť napríklad z hľadiska materiálno-technickej vybavenosti škôl (existencia školských laboratórií, dostupnosť digitálnych technológií a ďalších didaktických prostriedkov) alebo s ohľadom na kompetencie učiteľov realizovať výučbu vedúcu k rozvoju bádateľských zručností žiakov (Dluhošová, 2004; Hew & Brush, 2007; Rahayu et al., 2022; Šterbáková, 2014). Táto situácia môže významnou mierou determinovať dosiahnutú úroveň bádateľských zručností žiakov.

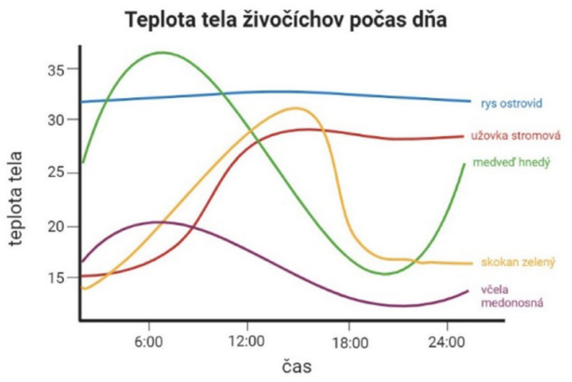
## 4.2 Výskumný nástroj

Za účelom zistenia aktuálnej úrovne bádateľských zručností žiakov existuje už niekoľko dostupných výskumných nástrojov (napr. Burns et al., 1985; Gormally et al., 2012; Ješková et al., 2016b; Wenning, 2007; Wenning, 2006 a pod.). Tieto výskumné nástroje sú v prevažnej miere určené pre starších žiakov či študentov, a z toho dôvodu ich nie je jednoduché prevziať a použiť na meranie zručností žiakov šiesteho ročníka nižšieho sekundárneho vzdelávania. Pre zistenie úrovne vybraných bádateľských zručností žiakov sme skonštruovali vlastný výskumný nástroj v podobe testu s uzavretými položkami (obr. 1). Každá položka obsahovala jednu správnu možnosť a štyri ponúkané distraktory, čím sa znižuje pravdepodobnosť náhodného označenia (uhádnutia) správnej odpovede (Farhady & Shakery, 2000; Hassan & Hod, 2017; Woodford & Bancroft, 2004). Test pozostával z celkovo 14 uzatvorených položiek vsadených do kontextu biológie, pričom každá zručnosť bola v teste meraná prostredníctvom dvoch položiek. Pri výbere konkrétnych bádateľských zručností sme vychádzali z prác Wenninga (2010; 2005), pričom sme vymedzili sedem bádateľských zručností, ktorými by mali žiaci šiesteho ročníka základných škôl disponovať, a to: zručnosť formulovať predpovede, zručnosť identifikovať premenné, zručnosť identifikovať vzťah medzi premennými na základe údajov z grafu, zručnosť identifikovať vzťah medzi premennými na základe údajov z tabuľky, zručnosť zaznamenávať výsledky pozorovania a merania, zručnosť transformovať výsledky do štandardných foriem a zručnosť formulovať záver. Na administráciu testu sme stanovili 45 minút.

Validitu výskumného nástroja sme zabezpečili expertným posúdením troch odborníkov z oblasti didaktiky (Heale & Twycross, 2015). Reliabilitu testu sme určili prostredníctvom vzorca Kudera a Richardsona

**Úloha 5**

Vedci hľadali odpoveď na otázku: *Ako sa mení teplota tela v závislosti od prostredia u piatich vybraných živočíchoch počas dňa?* Na základe pokusu získali údaje, ktoré zaznamenali do grafu. Jedným zo skúmaných živočíchov bol aj skokan zelený. Ktorá z uvedených odpovedí obsahuje správnu formuláciu záveru pokusu s ohľadom na údaje v priloženom grafe?




a) Teplota tela skokana zeleného sa počas dňa nemení v závislosti od teploty prostredia, ide preto o živočícha s nestálou teplotou tela.

b) Teplota tela skokana zeleného sa počas dňa nemení v závislosti od teploty prostredia, ide preto o živočícha so stálou teplotou tela.

c) Teplota tela skokana zeleného sa počas dňa mení v závislosti od teploty prostredia, ide preto o živočícha so stálou teplotou tela.

d) Teplota tela skokana zeleného sa počas dňa mení v závislosti od teploty prostredia, ide preto o živočícha s nestálou teplotou tela.

e) Teplota okolitého prostredia nemá žiaden vplyv na teplotu tela skokana zeleného počas dňa.

**Obr. 1:** Ukážka testovej položky (položka 5)

č. 20, pretože jednotlivé položky v teste boli skórované dichotomicky. Zo zistenej hodnoty ($KR_{20} = 0{,}78$) vyplýva, že výskumný nástroj môžeme považovať za reliabilný (Jacob, 2017). Priemerná hodnota obťažnosti testových položiek $Q$ predstavovala 57,6 %, pričom čím je jeho hodnota vyššia, tým sú položky obťažnejšie (Kubiš et al., 2015). Priemerný index citlivosti položiek ULI (upper-lower index) dosiahol hodnotu 0,65, pričom čím je jeho hodnota bližšie k jednej, tým majú položky vyššiu diskriminačnú schopnosť (Noor, 2021). Jednotlivé hodnoty indexu obťažnosti a citlivosti sú uvedené v tab. 2.

**Tab. 2:** Hodnota indexu obťažnosti a citlivosti pre jednotlivé testové položky

| Položka | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Index citlivosti | 0,41 | 0,72 | 0,71 | 0,70 | 0,74 | 0,81 | 0,82 |
| Index obťažnosti $Q$ | 65,3 | 46,6 | 53,6 | 64,1 | 66,9 | 49,0 | 41,4 |

| Položka | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|
| Index citlivosti | 0,34 | 0,71 | 0,76 | 0,27 | 0,71 | 0,51 | 0,90 |
| Index obťažnosti $Q$ | 76,0 | 55,6 | 55,2 | 65,5 | 54,7 | 69,0 | 44,0 |

## 4.3 Analýza dát

Výsledky testu sme podrobili kvantitatívnej analýze, ktorá spočívala v stanovení základných opisných charakteristík testu (napr. aritmetický priemer, medián, modus, smerodajná odchýlka a pod.). Úspešnosť žiakov pri riešení testových položiek sme vypočítali ako podiel žiakov, ktorí uviedli správnu odpoveď k celkovému počtu testovaných žiakov (Prokša et al., 2008). Obdobne sme postupovali aj pri určovaní úspešnosti dosiahnutej v rámci jednotlivých bádateľských zručností. Shapiro-Wilkov test preukázal, že dáta nie sú normálne rozložené ($W = 0{,}93$, $p < 0{,}05$), a preto sme na ich analýzu použili neparametrické štatistické testy (Neideen & Brasel, 2007). Medzi takéto testy patrí Spearmanov korelačný koeficient, ktorý slúži na zistenie štatisticky významnej korelácie medzi dvomi poradovými premennými (Schober et al., 2018) a Mann Whitneyho (Wilcoxonov) test, ktorý sa používa na zistenie štatisticky významného rozdielu medzi mediánmi dvoch nezávislých výberov (Fagerland & Sandvik, 2009). Následne sme zisťovali aj veľkosť efektu $r$, ktorý predstavuje silu rozdielu medzi skupinami (Pallant, 2007).

# 5 Výsledky

## 5.1 Celková úroveň bádateľských zručností

Žiaci v administrovanom teste bádateľských zručností dosiahli priemerné skóre 5,93 bodov ($SD = 3{,}43$) z celkového počtu 14 bodov, čo predstavuje úspešnosť na úrovni 42,4 %. Medián bol na úrovni 6,0 a modus 1,0.

Najvyššiu úspešnosť žiakov sme zaznamenali pri zručnosti identifikovať vzťah medzi premennými na základe údajov v tabuľke (57,3 %), kde v oboch položkách (položka 7 a položka 14) mali žiaci vyvodiť správny vzťah medzi vymedzenou závisle a nezávisle premennou. O niečo nižšiu úspešnosť dosiahli žiaci pri samotnej zručnosti identifikovať premenné (48,9 %). V oboch položkách (položka 2 a položka 9) mali žiaci na základe navrhnutého postupu pokusu rozhodnúť, ktorý z uvedených faktorov má alebo nemá vplyv na jeho výsledok. Úspešnosť nad priemernou úspešnosťou riešenia testu sme zaznamenali aj pri zručnosti zaznamenávať výsledky pozorovania a merania (45,6 %), kde mali v položke 3 a v 10 posúdiť, či sa jedná o vhodný alebo nevhodný spôsob zaznamenávania údajov. Priemerné percentuálne skóre žiakov dosiahnuté v jednotlivých bádateľských zručnostiach uvádzame v tab. 3.

**Tab. 3:** Priemerné skóre žiakov dosiahnuté v teste

| zručnosť | položka | úspešnosť pre položku [%] | úspešnosť pre zručnosť [%] |
|---|---|---|---|
| zručnosť formulovať predpovede | 1<br>8 | 34,7<br>24,0 | 29,4 |
| zručnosť identifikovať premenné | 2<br>9 | 53,4<br>44,4 | 48,9 |
| zručnosť identifikovať vzťah medzi premennými na základe údajov z grafu | 6<br>13 | 51,0<br>31,0 | 41,0 |
| zručnosť identifikovať vzťah medzi premennými na základe údajov z tabuľky | 7<br>14 | 58,6<br>56,0 | 57,3 |
| zručnosť zaznamenávať výsledky pozorovania a merania | 3<br>10 | 46,4<br>44,8 | 45,6 |
| zručnosť transformovať výsledky do štandardných foriem | 4<br>11 | 35,9<br>34,5 | 35,2 |
| zručnosť formulovať záver | 5<br>12 | 33,1<br>45,3 | 39,2 |

Najnižšiu priemernú úspešnosť (tab. 3) dosiahli žiaci v zručnosti formulovať predpovede (29,4 %). V oboch položkách (položka 1 a položka 8) merajúcich túto zručnosť si mali žiaci vybrať správnu formuláciu predpovede spomedzi distraktorov, ktoré obsahovali navrhnutú výskumnú otázku, postup pozorovania, vysvetlenie pozorovania a pod. Nízke skóre úspešnosti žiakov naznačuje, že nedokážu odlíšiť predpoveď od iných výrokov a zároveň majú problém pochopiť kľúčové prvky správne naformulovanej predpovede.

Podpriemerné skóre žiaci získali aj v zručnosti transformovať výsledky do štandardných foriem. Priemerná úspešnosť žiakov bola na úrovni 35,2 %. Žiaci pri jednotlivých položkách (položka 4 a položka 11) mali problém s výberom správneho grafu, ktorý by najvhodnejšie reprezentoval grafické spracovanie vopred poskytnutých údajov. Problém žiakov s grafickým spracovaním dát sa ukázal aj pri zručnosti identifikovať vzťah medzi premennými na základe údajov z grafu, kde dosahovali rovnako podpriemernú úroveň (41,0 %).

V zručnosti formulovať záver dosiahli žiaci úspešnosť len 39,2 %. V oboch položkách merajúcich túto zručnosť (položka 5 a 12) mali zvoliť správne sformulovaný záver, ktorý by bolo možné vytvoriť na základe dostupných údajov. Analýzou získaných údajov sa ukázalo, že žiaci nedokážu dostatočne zovšeobecniť poskytnuté informácie a vyvodiť z nich relevantné závery.

Prostredníctvom Spearmanovho korelačného koeficientu sme zisťovali aj koreláciu medzi jednotlivými zručnosťami. Ukazuje sa, že medzi takmer všetkými vybranými zručnosťami existuje štatisticky významná korelácia (tab. 4), avšak ich hodnoty sú slabé až stredne silné (Schober et al., 2018). Medzi zručnosťami formulovať predpovede a transformovať výsledky do štandardných foriem sme nezistili žiadnu štatisticky významnú koreláciu.

**Tab. 4:** Spearmanov korelačný koeficient medzi jednotlivými bádateľskými zručnosťami

| zručnosť | formulovať predpovede | identifikovať premenné | identifikovať vzťah medzi premennými na základe údajov z grafu | identifikovať vzťah medzi premennými na základe údajov z tabuľky | zaznamenávať výsledky pozorovania a merania | transformovať výsledky do štandardných foriem | formulovať záver |
|---|---|---|---|---|---|---|---|
| formulovať predpovede | × | 0,12* | 0,23* | 0,19* | 0,25* | 0,05 | 0,25* |
| identifikovať premenné | | × | 0,43* | 0,57* | 0,47* | 0,07* | 0,39* |
| identifikovať vzťah medzi premennými na základe údajov z grafu | | | × | 0,50* | 0,38* | 0,28* | 0,37* |
| identifikovať vzťah medzi premennými na základe údajov z tabuľky | | | | × | 0,55* | 0,06* | 0,49* |
| zaznamenávať výsledky pozorovania a merania | | | | | × | 0,05* | 0,49* |
| transformovať výsledky do štandardných foriem | | | | | | × | 0,08* |
| formulovať záver | | | | | | | × |
| Priemerné skóre za zručnosť | 0,29 | 0,49 | 0,41 | 0,57 | 0,45 | 0,35 | 0,39 |
| *SD* | 0,48 | 0,49 | 0,50 | 0,48 | 0,49 | 0,43 | 0,42 |

* hladina významnosti 95 % ($p < 0{,}05$)

## 5.2 Úroveň bádateľských zručností žiakov v závislosti od pohlavia

Porovnaním získaných údajov vzhľadom na pohlavie sme zistili, že chlapci ($x = 5{,}88$; $SD = 3{,}40$) aj dievčatá ($x = 5{,}98$; $SD = 3{,}46$) dosiahli v teste porovnateľné priemerné skóre. Mann-Whitneyho (Wilcoxonov) test nepotvrdil štatisticky významný rozdiel ($W = -1\,375{,}5$; $p = 0{,}72$) na hladine významnosti 95 % medzi úrovňou bádateľských zručností žiakov a pohlavím. Následne sme zisťovali, či existuje štatisticky významný rozdiel medzi úrovňou konkrétnych zručností a pohlavím žiakov. Štatisticky významný rozdiel na hladine významnosti 95 % v prospech dievčat sme zistili iba pri zručnosti zaznamenávať výsledky pozorovania a merania (tab. 5), pričom veľkosť účinku ($r = -0{,}1$) môžeme považovať za triviálnu (Cohen, 1988).

**Tab. 5:** Výsledky štatistickej analýzy úrovne jednotlivých zručností vzhľadom na pohlavie

| | | chlapci | | dievčatá | | | |
|---|---|---|---|---|---|---|---|
| **zručnosť** | **položka** | **úspešnosť položky (%)** | **úspešnosť zručnosti (%)** | **úspešnosť položky (%)** | **úspešnosť zručnosti (%)** | **W** | **p-value** |
| zručnosť formulovať predpovede | 1 | 34,5 | 30,3 | 34,8 | 28,4 | −7 758,0 | 0,36 |
| | 8 | 26,2 | | 22,1 | | | |
| zručnosť identifikovať premenné | 2 | 52,0 | 47,9 | 54,7 | 49,7 | 7 110,0 | 0,45 |
| | 9 | 43,9 | | 44,8 | | | |
| zručnosť identifikovať vzťah medzi premennými na základe údajov z grafu | 6 | 48,9 | 39,6 | 52,7 | 42,2 | 10 305,0 | 0,26 |
| | 13 | 30,3 | | 31,6 | | | |
| zručnosť identifikovať vzťah medzi premennými na základe údajov z tabuľky | 7 | 58,1 | 57,3 | 58,9 | 57,2 | −252,0 | 0,98 |
| | 14 | 56,5 | | 55,5 | | | |
| zručnosť zaznamenávať výsledky pozorovania a merania* | 3 | 42,5 | 42,6 | 49,7 | 48,1 | 21 825,0 | 0,02 |
| | 10 | 42,7 | | 46,5 | | | |
| zručnosť transformovať výsledky do štandardných foriem | 4 | 39,9 | 37,1 | 32,3 | 33,4 | −14 553, 0 | 0,11 |
| | 11 | 34,3 | | 34,6 | | | |
| zručnosť formulovať záver | 5 | 33,5 | 38,9 | 32,6 | 39,6 | 2 538,0 | 0,78 |
| | 12 | 44,2 | | 43,4 | | | |

*hladina významnosti 95 % ($p < 0{,}05$)

# 6 Diskusia

Kireš a Jurková (2021) považujú za dobre rozvinutú takú zručnosť, v ktorej žiaci dosiahli úspešnosť nad 75 %. Výsledky výskumu žiakov šiesteho ročníka základných škôl na Slovensku však odhalili, že disponujú výrazne nižšou úrovňou bádateľských zručností (42,4 %), čo považujeme za neuspokojivý výsledok. Analýza jednotlivých zručností navyše preukázala, že aj keď existujú štatisticky významné korelácie medzi mnohými skúmanými zručnosťami, tieto korelácie dosahujú častokrát nízke hodnoty. Medzi bádateľskými zručnosťami by mali byť silné korelácie nakoľko je to vzájomne sa prelínajúci súbor zručností, ktoré sú navzájom od seba závislé a ktoré by si mali žiaci v konečnom dôsledku osvojiť na takej úrovni, aby boli schopní nájsť riešenie rôznych problémov prostredníctvom realizácie bádateľských aktivít (Kruit et al., 2018).

Pri detailnejšej analýze výsledkov testu sme zaznamenali výrazne nižšiu percentuálnu úspešnosť žiakov pri zručnosti formulovať predpovede ako vo svojich výskumoch zistili autori Šmida a Čipková (2021), Kireš a Jurková (2021) či Öztürk et al. (2010). Žiaci si často zamieňali predpoveď s inými výrokmi, pričom práve schopnosť odlíšiť predpoveď od všeobecných tvrdení či otázky považujú Krišková a Kireš (2017) za základnú a nevyhnutnú etapu rozvoja tejto zručnosti. Žiaci mali tiež problém so zručnosťou transformovať výsledky do štandardných foriem aj napriek tomu, že by mali byť na vyučovaní vedení k systematickému zaznamenávaniu údajov (Etkina et al., 2006; Giammatteo & Obaya, 2018; Liew et al., 2019) do grafov či tabuliek, ktoré slúžia na sprehľadnenie výsledkov a zachytenie trendov medzi údajmi (Orolínová & Kotuľáková, 2014). Rovnako tak žiaci mali problém s identifikáciou vzťahu medzi premennými na základe údajov z grafu, čo naznačuje, že sa u nich vyskytujú nie len ťažkosti s transformáciou údajov, ale aj s ich následnou analýzou, ktorá je nevyhnutná pri hľadaní vzťahu medzi premennými. Beaumont-Walters a Soyibo (2001) vo svojom výskume tiež poukazujú na skutočnosť, že žiaci majú ťažkosti s prácou s grafmi, pretože si vyžaduje pokročilú schopnosť rozpoznávať vzťahy medzi údajmi. Glazer (2011) však prácu žiakov s grafmi považuje za esenciálny prvok pri rozvíjaní prírodovednej gramotnosti, a preto je potrebné mu vo vyučovaní prírodovedných predmetov venovať dostatok pozornosti.

Problémy žiakov sme zaznamenali aj pri zručnosti formulovať záver, na čo poukazuje aj NRC (1998). Žiaci v teste nedokázali dostatočne zovšeobecniť poskytnuté relevantné údaje a transformovať ich do vhodne naformulovaného záveru. Aj podľa Orolínovej a Kotuľákovej (2014) žiaci pri tvorbe záverov robia často chybu v tom, že ignorujú relevantné dôkazy, nevedia vytvoriť logické prepojenie medzi dôkazmi a vysvetleniami a vytvárajú závery, ktoré nie je možné na základe ich údajov vyvodzovať. Lati et al. (2012) uvádzajú, že problémy žiakov s formuláciou záverov môžu byť spôsobené tým, že sa jedná o časovo náročnú činnosť, ktorá je v rámci realizácie praktických aktivít vo vyučovaní často vynechávaná. Učiteľ by si však na túto činnosť mal vyhradiť dostatok času, pretože najskôr musí žiakom ukázať, čo má záver obsahovať a naučiť ich, ktoré údaje sú dostatočne relevantné na to, aby ich mohli zovšeobecniť a aby následne mohli slúžiť na podporu vyvodených záverov (Nurdin et al., 2019).

Analýza získaných údajov zároveň nepreukázala štatisticky významné rozdiely v celkovej úrovni bádateľských zručností žiakov vzhľadom na pohlavie, čo sa potvrdilo aj pri výskumoch realizovaných u starších žiakov (napr. Čipková et al., 2020; Ong et al., 2015; Öztürk et al., 2010; Šmida & Čipková, 2021). Túto skutočnosť považujeme za pozitívny odklon od klasických rodových stereotypov vo vyučovaní prírodovedných predmetov, kedy chlapci bývajú pri riešení podobných testov úspešnejší ako dievčatá (napr. Ješková et al., 2021; Ješková et al., 2016b; Nosálová, 2022), pretože obe pohlavia by mali dostávať rovnakú príležitosť pri rozvoji svojich zručností (Rao, 2008). Štatisticky významné rozdiely sme zaznamenali iba pri zručnosti zaznamenávať výsledky pozorovania a merania, ktoré boli v prospech dievčat, čo podporuje tvrdenie Delena a Kesercioğlua (2012), že zatiaľ čo sa chlapci aktívnejšie zapájajú do realizácie experimentov, dievčatá skôr vystupujú v úlohe pozorovateliek a precíznejšie si zaznamenávajú získané údaje do protokolu. Veľkosť účinku medzi oboma skupinami je však triviálna (Cohen, 1988), a preto je nevyhnutné detailnejšie preskúmať, aký vplyv môže mať pohlavie žiakov na uvedenú bádateľskú zručnosť.

# 7 Záver

Problematika bádania a rozvoja zručností žiakov vo vyučovaní prírodných vied je na Slovensku diskutovanou témou. Napriek tomu, že inovovaný Štátny vzdelávací program (ŠPÚ, 2015) presadzuje spôsob vzdelávania, ktorý umožňuje žiakom bádať a objavovať, realizované výskumy dlhodobo poukazujú na nízku úroveň bádateľských zručností žiakov základných škôl (napr. Kireš & Jurková, 2021; Šmida & Čipková, 2021; Zheng et al., 2022), čo potvrdili aj výsledky nášho výskumu u žiakov 6. ročníka nižšieho sekundárneho vzdelávania. Nízka úroveň zručností bola zaznamenaná aj na úrovni stredných škôl (napr. Hodosyova et al., 2015; Ješková et al., 2021; Ješková et al., 2018; Ješková et al., 2016a; Ješková et al., 2016b) a vysokých škôl (Čipková & Fuchs, 2020; Čipková & Karolčík, 2018; Fehér et al., 2020).

Výsledky výskumu zároveň preukázali, že žiaci majú najväčší problém so zručnosťami spojenými s formuláciou predpovede, s transformovaním výsledkov do štandardných foriem a s formulovaním záverov. Nízka celková úroveň osvojených bádateľských zručností žiakov 6. ročníka je alarmujúca, pretože práve skúmané zručnosti podľa Wenninga (2005) patria do skupiny elementárnych a základných bádateľských zručností. Osvojenie si tejto skupiny zručností predstavuje základný predpoklad pre rozvoj pokročilých bádateľských zručností, čo v konečnom dôsledku predstavuje jeden zo základných cieľov prírodovedného vzdelávania (NRC, 2000; Wenning, 2010). Kvalitatívny rozvoj týchto zručností je možné preto zabezpečiť prostredníctvom cielenej a systematickej implementácie bádateľských aktivít do vzdelávacieho procesu, čo umožní žiakom nadobúdať cenné skúsenosti spojené s prácou vedcov (Sparks & Deane, 2015), zvýšiť ich záujem a motiváciu k vlastnému vzdelávaniu v oblasti prírodných vied, pretaviť konštruktivistickú teóriu priamo do praxe (Justice et al., 2007) a zvýšiť tak úroveň prírodovednej gramotnosti žiakov.

## 8 Limity výskumu

Limitom výskumu je použitie testu s uzavretými položkami, ktoré umožňujú zistiť iba deklaratívnu úroveň bádateľských zručností a nie ich hĺbku. Test navyše obsahoval len 14 položiek, pričom každá zručnosť bola meraná prostredníctvom dvoch položiek, čo neumožňuje komplexnejšie posúdenie úrovne zručností žiakov. Počet testových položiek sme však navrhli s prihliadnutím na vývinovú úroveň žiakov šiesteho ročníka základných škôl.

## Literatúra


Balogová, B., & Ješková, Z. (2016). Mapovanie bádateľských zručností žiakov stredných škôl. *Biológia, ekológia, chémia*, *20*(3). https://bit.ly/3MwtbaM

Beaumont-Walters, Y., & Soyibo, K. (2001). An analysis of high school students' performance on five integrated science process skills. *Research in Science & Technological Education*, *19*(2), 133–145. https://doi.org/10.1080/02635140120087687

Bellová, R., Melicherčíková, D., & Tomčík, P. (2018). Possible reasons for low scientific literacy of Slovak students in some natural science subjects. *Research in Science & Technological Education*, *36*(2), 226–242. https://doi.org/10.1080/02635143.2017.1367656

Burns, J. C., Okey, J. R., & Wise, K. C. (1985). Development of an integrated process skill test: TIPS II. *Journal of Research in Science Teaching*, *22*(2), 169–177. https://doi.org/10.1002/tea.3660220208

Cheung, D. (2007). Facilitating chemistry teachers to implement inquiry-based laboratory work. *International Journal of Science and Mathematics Education*, *6*(1), 107–130. https://doi.org/10.1007/s10763-007-9102-y

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillsdale, NJ: Erlbaum.

Colburn, A. (2000). An inquiry primer. *Science Scope*, *23*(6), 42–44. https://www.studentachievement.org/wp-content/uploads/An-Inquiry-Primer-1.pdf

Čepičková, I. B. (2013). *Didaktika přírodovědného základu.* Univerzita J. E. Purkyně.

Čipková, E., & Fuchs, M. (2020). Hodnotenie vybraných bádateľských zručností študentov učiteľstva biológie. *Scientia in educatione*, *11*(2), 2–13. https://doi.org/10.14712/18047106.1884

Čipková, E., & Karolčík, Š. (2018). Assessing of scientific inquiry skills achieved by future biology teachers. *Chemistry-Didactics-Ecology-Metrology*, 23. https://doi.org/10.1515/cdem-2018-0004

Čipková, E., Karolčík, Š., & Scholzová, L. (2020). Are secondary school graduates prepared for the studies of natural sciences? – evaluation and analysis of the result of scientific literacy levels achieved by secondary school graduates. *Research in Science & Technological Education*, *38*(2), 146–167. https://doi.org/10.1080/02635143.2019.1599846

DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *37*(6), 582–601. https://doi.org/10.1002/1098-2736(200008)37:6%3C582::AID-TEA5%3E3.0.CO;2-L

Delen, İ., & Kesercioğlu, T. (2012). How middle school students' science process skills affected by Turkey's national curriculum change? *Journal of Turkish Science Education*, *9*(4), 3–9. http://www.tused.org/index.php/tused/article/view/465/399

Deters, K. M. (2005). Student opinions regarding inquiry-based labs. *Journal of Chemical Education*, *82*(8), 1 178–1 180.

Dluhošová, A. (2004). Rómske deti v školstve z pohľadu školskej inšpekcie. *Rómske deti v slovenskom školstve*, 42. https://bit.ly/3EITHfo

Ergül, R., Şımşeklı, Y., Çaliş, S., Özdılek, Z., Göçmençelebı, Ş., & Şanli, M. (2011). The effects of inquiry-based science teaching on elementary school students'science process skills and science attitudes. *Bulgarian Journal of Science & Education Policy*, *5*(1). http://see-articles.ceon.rs/data/pdf/1313-1958/2011/1313-19581101048E.pdf

Etkina, E., Van Heuvelen, A., White-Brahmia, S., Brookes, D. T., Gentile, M., Murthy, S., Rosengrant, D., & Warren, A. (2006). Scientific abilities and their assessment. *Physical Review Special Topics-Physics Education Research*, *2*(2). https://journals.aps.org/prper/abstract/10.1103/PhysRevSTPER.2.020103

Fagerland, M. W., & Sandvik, L. (2009). The wilcoxon–mann–whitney test under scrutiny. *Statistics in Medicine*, *28*(10), 1 487–1 497. https://doi.org/10.1002/sim.3561

Farhady, H., & Shakery, S. (2000). Number of options and economy of multiple-choice tests. *Roshd Foreign Language Teaching Journal*, *14*(1), 132–141. https://bit.ly/3VsvRuc

Fehér, Z., Jaruska, L., Szarka, K., & Tóth, P. (2020). Assessing the condition of scientific inquiry skills in teacher education using physical and chemical tasks. In *ICERI2020 Proceedings* (pp. 5 416–5 424). IATED. https://library.iated.org/view/FEHER2020ASS

Firman, H. (2007). *Laporan Analisis Lietarasi Sains Berdasarkan Hasil PISA Nasional Tahun 2006.* Pusat Penilaian Pendidikan Balitbang Depdiknas, Jakarta.

Fradd, S. H., Lee, O., Sutman, F. X., & Saxton, M. K. (2001). Promoting! Science literacy with English language learners through instructional materials development: A case study. *Bilingual Research Journal*, *25*(4), 417–439. https://doi.org/10.1080/15235882.2001.11074464

Glazer, N. (2011). Challenges with graph interpretation: a review of the literature. *Studies in Science Education*, *47*(2), 183–210. https://doi.org/10.1080/03057267.2011.605307

Giammatteo, L., & Obaya, A. V. (2018). Assessing chemistry laboratory skills through a competency-based approach in high school chemistry course. *Science Education International*, *29*(2), 103–109. https://eric.ed.gov/?id=EJ1184773

Gormally, C., Brickman, P., & Lutz, M. (2012). Developing a test of scientific literacy skills (TOSLS): Measuring undergraduates' evaluation of scientific information and arguments. *CBE-Life Sciences Education*, *11*(4), 364–377. https://doi.org/10.1187/cbe.12-03-0026

Guevara, C. A. (2015). Science process skills development through innovations in science teaching. *Research Journal of Educational Sciences*, *3*(2), 6–10. http://www.isca.in/EDU_SCI/Archive/v3/i2/2.ISCA-RJEduS-2015-003.php

Hairida, H. (2016). The effectiveness using inquiry based natural science module with authentic assessment to improve the critical thinking and inquiry skills of junior high school students. *Jurnal Pendidikan IPA Indonesia*, *5*(2), 209–215.

Harrison, C. (2014). Assessment of inquiry skills in the SAILS project. *Science Education International*, *25*(1), 112–122. https://eric.ed.gov/?id=EJ1022890

Harlen, W. (2001). The assessment of scientific literacy in the OECD/PISA project. *Studies in Science Education*, *36*(1), 79–103. https://doi.org/10.1080/03057260108560168

Harlen, W. (2013). *Assessment & inquiry-based science education: Issues in policy and practice.* Science Education Program of IAP. https://bit.ly/3rVNmFD

Harlen, W. (2014). Helping children's development of inquiry skills. *Inquiry in Primary Science Education*, *1*(1), 5–19. https://bit.ly/3TiBhWM

Hassan, S., & Hod, R. (2017). Use of item analysis to improve the quality of single best answer multiple choice question in summative assessment of undergraduate medical students in Malaysia. *Education in Medicine Journal*, *9*(3), 33–43. https://doi.org/10.21315/eimj2017.9.3.4

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, *18*(3), 66–67. https://doi.org/10.1136/eb-2015-102129

Hew, K. F., & Brush, T. (2007). Integrating technology into K-12 teaching and learning: Current knowledge gaps and recommendations for future research. *Educational Technology Research and Development*, 55(3), 223–252. https://link.springer.com/article/10.1007/s11423-006-9022-5

Hodosyova, M., Útla, J., Vnukova, P., & Lapitkova, V. (2015). The development of science process skills in physics education. *Procedia-Social and Behavioral Sciences*, 186, 982–989. https://doi.org/10.1016/j.sbspro.2015.04.184

Holbrook, J., & Rannikmae, M. (2009). The meaning of scientific literacy. *International Journal of Environmental and Science Education*, *4*(3), 275–288. https://files.eric.ed.gov/fulltext/EJ884397.pdf

Jacob, J. (2017). Reliability: How? When? What? *International Journal of Advances in Nursing Management*, *5*(4), 372–374. https://bit.ly/3T1OEef

Ješková, Z., Balogová, B., & Kireš, M. (2018). Assessing inquiry skills of upper secondary school students. In *Journal of Physics: Conference Series* (vol. 1076, no. 1). IOP Publishing. https://iopscience.iop.org/article/10.1088/1742-6596/1076/1/012022/meta

Ješková, Z., Jurková, V., Lukáč, S., Šnajder, Ľ., & Guniš, J. (2021). Development of inquiry skills at upper secondary level. *Journal of Physics, Conf. Series*, *1929*, 1–11. https://iopscience.iop.org/article/10.1088/1742-6596/1929/1/012029/meta

Ješková, Z., Lukáč, S., Hančová, M., Šnajder, Ľ., Guniš, J., Balogová, B., & Kireš, M. (2016a). Efficacy of inquiry-based learning in mathematics, physics and informatics in relation to the development of students inquiry skills. *Journal of Baltic Science Education*, *15*(5), 559–574. https://bit.ly/3g32K0h

Ješková, Z., Lukáč, S., Šnajder, Ľ., Guniš, J., Balogová, B., & Kireš, M. (2016b). Hodnotenie bádateľských zručností žiakov gymnáziá. *Scientia in educatione*, *7*(2), 48–70. https://doi.org/10.14712/18047106.350

Justice, C., Rice, J., Warry, W., Inglis, S., Miller, S., & Sammon, S. (2007). Inquiry in higher education: Reflections and directions on course design and teaching methods. *Innovative Higher Education*, *31*(4), 201–214. https://link.springer.com/article/10.1007/s10755-006-9021-9

Kireš, M., Ješková, Z., Ganajová, M., & Kimáková, K. (2016). *Bádateľské aktivity v prírodovednom vzdelávaní.* Štátny pedagogický ústav. https://https://bit.ly/3VvR0n6

Kireš, M., & Jurková, V. (2021). Development of inquiry skills at lower secondary school level. *Journal of Physics: Conference Series*, *1929*(1). https://iopscience.iop.org/article/10.1088/1742-6596/1929/1/012028/meta

Kotuľáková, K. (2020). Prírodovedná gramotnosť a kritické zhodnotenie mediálnych výstupov. *Biológia, Ekológia, Chémia*, *24*(1). https://bit.ly/3T3TLdY

Krišková, K., & Kireš, M. (2017). Making predictions skill level analysis. In *AIP Conference Proceedings* (vol. 1804, no. 1). AIP Publishing LLC. https://doi.org/10.1063/1.4974393

Kruit, P. M., Oostdam, R. J., van den Berg, E., & Schuitema, J. A. (2018). Assessing students' ability in performing scientific inquiry: instruments for measuring science skills in primary education. *Research in Science & Technological Education*, *36*(4), 413–439.

Kubiš, T., Demkanin, P., Hajdúk, M., Hanuljaková, H., Lapitka, M., & Malčík, M. (2015). *Metodika tvorby testových úloh a testov.* Národný ústav certifikovaných meraní vzdelávania.

Lati, W., Supasorn, S., & Promarak, V. (2012). Enhancement of learning achievement and integrated science process skills using science inquiry learning activities of chemical reaction rates. *Procedia-Social and Behavioral Sciences*, 46, 4 471–4 475. https://doi.org/10.1016/j.sbspro.2012.06.279

Liew, S. S., Lim, H. L., Saleh, S., & Ong, S. L. (2019). Development of scoring rubrics to assess physics practical skills. *EURASIA Journal of Mathematics, Science and Technology Education*, *15*(4), 1–14. https://doi.org/10.29333/ejmste/103074

Llewellyn, D. (2013). *Teaching high school science through inquiry and argumentation – Second edition.* Corwin, A SAGE Company.

Liu, X. (2013). Expanding notions of scientific literacy: A reconceptualization of aims of science education in the knowledge society. In *Science education for diversity* (pp. 23–39). Springer, Dordrecht. https://link.springer.com/chapter/10.1007/978-94-007-4563-6_2

Lou, Y., Blanchard, P., & Kennedy, E. (2015). Development and validation of a science inquiry skills assessment. *Journal of Geoscience Education*, *63*(1), 73–85. https://doi.org/10.5408/14-028.1

Miklovičová, J., & Valovič, J. (2019). *Národná správa PISA 2018.* Bratislava: Národný ústav certifikovaných meraní vzdelávania.

Millar, R., & Driver, R. (1987). Beyond Processes, *Studies in Science Education.* https://doi.org/10.1080/03057268708559938

Minárechová, M. (2014). História induktívneho prístupu v prírodovednom vzdelávaní v USA a jeho súčasná reflexia na Slovensku. *Scientia in educatione*, *5*(1), 2–19. https://doi.org/10.14712/18047106.94

Miškovičová Hunčíková, I., & Ušáková, K. (2009). Experimentálne overovanie alternatívneho obsahu biológie na gymnáziách. In *New trends in the didactic training of teachers* (pp. 19–24). EDUCO.

Mukti, W. R., Yuliskurniawati, I. D., Noviyanti, N. I., Mahanal, S., & Zubaidah, S. (2019). A survey of high school students' scientific literacy skills in different gender. *Journal of Physics: Conference Series*, *1241*(1). https://iopscience.iop.org/article/10.1088/1742-6596/1241/1/012043/meta

Mullis, I. V., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science.* TIMSS & PIRLS International Study Center, Boston College.

Mulyeni, T., Jamaris, M., & Supriyati, Y. (2019). Improving basic science process skills through inquiry-based approach in learning science for early elementary students. *Journal of Turkish Science Education*, *16*(2), 187–201.

Neideen, T., & Brasel, K. (2007). Understanding statistical tests. *Journal of Surgical Education*, *64*(2), 93–96. https://doi.org/10.1016/j.jsurg.2007.02.001

Nejedlý, A., & Vojíř, K. (2022). How do students formulate a research question and conclusions in science research? In *Project-Based Education And Other Student-Activation Strategies And Issues In Science Education XIX.* (pp. 29–38). PedF UK.

Noor, A. M. (2021). Evaluating multple choice questions from engineering statistics assessment. *International Journal of Education and Pedagogy*, *3*(4), 33–46. http://myjms.mohe.gov.my/index.php/ijeap

Nosálová, K. (2022). *Praktické cvičenia ako nástroj podpory bádateľských zručností.* PriF UK v Bratislave.

NRC. (1996). *National science education standards.* National Academies Press. https://bit.ly/3g8S40c

NRC. (1998). *The Canadian system of soil classification* (No. 1646). NRC Research Press.

NRC. (2000). *Inquiry and the national science education standards: A guide for teaching and learning.* National Academies Press. https://bit.ly/3yHeHil

Nurdin, K., Muh, H. S., & Muhammad, M. H. (2019). The implementation of inquiry-discovery learning. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, *7*(1), 164–175. https://doi.org/10.24256/ideas.v7i1.734

O'Connor, G., & Rosicka, C. (2020). *Science in the early years. Paper 2: Science inquiry skills.* Australian Council for Educational Research. https://bit.ly/3rTtLpA

Ong, E. T., Ramiah, P., Ruthven, K., Salleh, S. M., Yusuff, N. A. N., & Mokhsein, S. E. (2015). Acquisition of basic science process skills among Malaysian upper primary students. *Research in Education*, *94*(1), 88–101. https://doi.org/10.7227/RIE.0021

Opara, J. A., & Oguzor, N. S. (2011). Inquiry instructional method and the school science curriculum. *Current Research Journal of Social Sciences*, *3*(3), 188–198.

Orolínová, M., & Kotuľáková, K. (2014). *Rozvoj spôsobilostí vedeckej práce v podmienkach kontinuálneho vzdelávania učiteľov.* Typi Universitatis Tyrnaviensis.

Özgelen, S. (2012). Students' science process skills within a cognitive domain framework. *Eurasia Journal of Mathematics, Science and Technology Education*, *8*(4), 283–292. https://bit.ly/3T3eyhL

ÖztÜrk, N., Tezel, Ö., & Acat, M. B. (2010). Science process skills levels of primary school seventh grade students in science and technology lesson. *Journal of Turkish Science Education*, *7*(3), 15–28. http://tused.org/index.php/tused/article/view/520

Pallant, J. (2007). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS.* Routledge.

Prokša, M., Held, Ľ., Haláková, Z., Tóthová, A., Orolínová, M., Urbanová, A., & Žoldošová, K. (2008). *Metodológia pedagogického výskumu a jeho aplikácia v didaktikách prírodných vied.* Bratislava: Univerzita Komenského.

Rahayu, U., Sekarwinahyu, M., & Sapriati, A. (2022). The inquiry skills of teachers in elementary school. *Jurnal Ilmiah Sekolah Dasar*, *6*(2). https://doi.org/10.23887/jisd.v6i2.46909

Rao, D. B. (2008). *Science process skills of school students.* Discovery Publishing House.

Rezba, R. J., Sprague, C., & Fiel, R. (2003). *Learning and assessing science process skills.* Kendall Hunt.

Shahali, E. H. M., & Halim, L. (2010). Development and validation of a test of integrated science process skills. *Procedia-Social and Behavioral Sciences*, *9*, 142–146. https://doi.org/10.1016/j.sbspro.2010.12.127

Schleicher, A. (2019). PISA 2018: Insights and interpretations. OECD Publishing.

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5), 1 763–1 768. https://doi.org/10.1213/ANE.0000000000002864

Song, Y. (2016). "We found the 'black spots' on campus on our own": Development of inquiry skills in primary science learning with BYOD (Bring Your Own Device). *Interactive Learning Environments*, *24*(2), 291–305. https://doi.org/10.1080/10494820.2015.1113707

Sparks, J. R., & Deane, P. (2015). Cognitively based assessment of research and inquiry skills: Defining a key practice in the English language arts. ETS *Research Report Series*, *2015*(2), 1–55. https://doi.org/10.1002/ets2.12082

Stone, E. M. (2014). Guiding students to develop an understanding of scientific inquiry: A science skills approach to instruction and assessment. *CBE — Life Sciences Education*, *13*(1), 90–101. https://doi.org/10.1187/cbe-12-11-0198

Suryawati, E., & Osman, K. (2017). Contextual learning: Innovative approach towards the development of students' scientific attitude and natural science performance. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(1), 61–76. https://doi.org/10.12973/ejmste/79329

Šmida, D., & Čipková, E. (2021). Inquiry skills of primary school pupils in Slovakia. In *4th ICTLE, Proceedings of The 4th International Conference on Teaching, Learning and Education* (pp. 84–97). Diamond Scientific Publishing. https://www.dpublication.com/wp-content/uploads/2021/08/31-6641.pdf

ŠPÚ. (2015). *Inovovaný ŠVP pre 2. stupeň ZŠ.* https://bit.ly/3EJkXu6

Šterbáková, K. (2014). Nové technológie – interaktívna tabuľa SMART Board vo vyučovaní fyziky. *Edukacja – Technika – Informatyka*, *5*(2), 181–186. https://bit.ly/3S4u1g9

Tan, A. (1996). The way forward for Penang: Growth or development? In *Penang Economic Seminar*, Penang, May.

Temiz, B. (2020). Assessing skills of identifying variables and formulating hypotheses using scenario-based multiple-choice questions. *International Journal of Assessment Tools in Education*, *7*(1), 1–17. https://doi.org/10.21449/ijate.561895

TIMSS (2019). Prvé výsledky medzinárodného výskumu vedomostí a zručností žiakov 4. ročníka ZŠ v matematike a prírodných vedách. https://www2.nucem.sk/dl/4840/Prv%C3%A9%20v%C3%BDsledky%20Slovenska%20v%20%C5%A1t%C3%BAdii%20TIMSS%202019.pdf

Topalsan, A. K. (2020). Development of scientific inquiry skills of science teaching through argument-focused virtual laboratory applications. *Journal of Baltic Science Education*, *19*(4), 628–646. https://eric.ed.gov/?id=EJ1264527

Tosun, C. (2019). Scientific process skills test development within the topic "Matter and its Nature" and the predictive effect of different variables on 7th and 8th grade students' scientific process skill levels. *Chemistry Education Research and Practice*, *20*(1), 160–174. 10.1039/C8RP00071A

Wang, J., Guo, D., & Jou, M. (2015). A study on the effects of model-based inquiry pedagogy on students' inquiry skills in a virtual physics lab. *Computers in Human Behavior*, 49, 658–669. https://doi.org/10.1016/j.chb.2015.01.043

Wenning, C. J. (2005). Levels of inquiry: Hierarchies of pedagogical practices and inquiry processes. *Journal of Physics Teacher Education Online*, *2*(3), 3–12. http://www2.phy.ilstu.edu/pte/publications/levels_of_inquiry.pdf

Wenning, C. J. (2006). Assessing nature-of-science literacy as one component of scientific literacy. *Journal of Physics Teacher Education Online*, *3*(4), 3–14. https://www.phy.ilstu.edu/jpteo

Wenning, C. J. (2007). Assessing inquiry skills as a component of scientific literacy. *Journal of Physics Teacher Education Online*, *4*(2), 21–24. http://www2.phy.ilstu.edu/pte/publications/assessing_ScInq.pdf

Wenning, C. J. (2010). Levels of inquiry: Using inquiry spectrum learning sequences to teach science. *Journal of Physics Teacher Education Online*, *5*(3), 11–20. http://www.phy.ilstu.edu/pte/publications/learning_sequences.pdf

Windschitl, M. (2000). Supporting the development of science inquiry skills with special classes of software. *Educational Technology Research and Development*, *48*(2), 81–95. https://link.springer.com/article/10.1007/BF02313402

Woodford, K., & Bancroft, P. (2004). Using multiple choice questions effectively in information technology education. *21st Annual ASCILITE Conference 2004*, *4*, 948–955. https://www.ascilite.org/conferences/perth04/procs/pdf/woodford.pdf

Zeidan, A. H., & Jayosi, M. R. (2015). Science process skills and attitudes toward science among Palestinian secondary school students. *World Journal of Education*, *5*(1), 13–24. https://doi.org/10.5430/wje.v5n1p13

Zheng, Y., Yu, S., Zhang, M., Wang, J., Yang, X., Zheng, S., & Ping, X. (2022). Research on performance assessment of students' inquiry skills in China's elementary schools: A video analysis of Beijing discovering science around us. *Research in Science & Technological Education*, 1–27. https://doi.org/10.1080/02635143.2022.2126973

# Secondary and university students' understanding of independence and conditional probability

František Mošna[1]

[1] Pedagogická fakulta UK, Magdalény Rettigové 4, 116 39 Praha 1, Czech Republic; frantisek.mosna@pedf.cuni.cz

The primary objective of the study was to investigate the perceptions of Czech secondary and university students regarding independence in tasks involving multiple repetitions of random events and their understanding of conditional probability. The study employed a sample of 43 students, ranging in age from 15 to 23, who engaged in think-aloud interviews. The selection of eight tasks was based on existing literature. A qualitative analysis of the interview transcripts established that students encountered difficulties comprehending the concepts of independence and conditional probability, irrespective of whether they had previously undertaken a university course on probability. Notably, certain misconceptions about independence only surfaced in more challenging tasks, wherein students relied more on their intuition than their acquired knowledge. The misconceptions primarily manifested when describing the random space. The research findings have significant educational implications, which are discussed.

Key words:
independence,
conditional probability,
intuitive perceptions,
teaching of probability.

Received 1/2023
Revised 5/2023
Accepted 6/2023

## 1 Introduction

In today's scientific, economic, and technical fields, as well as in society and schools, stochastic and statistical models are of great importance. We encounter them in professional and daily life.

> We all rely on models to interpret everyday experiences. We interpret what we see in terms of mental models constructed on the basis of past experience and education. They are constructs that we use to understand the pattern of our experience. (Biarholomew quoted in Graham, 2006, p. 194)

Researchers agree that it is important that students come to university with the necessary level of understanding of concepts related to uncertainty and randomness (Nemirovsky et al., 2009). However, the importance of stochastic thinking is not always sufficiently appreciated. Studies on teaching probability show that students often have vague ideas about concepts related to uncertainty, randomness, independence and probability when moving from secondary school to university (Albert, 2003; Batanero, 2015; Evans, 2007; Fischbein, 1975; Fischbein & Schnarch, 1997; Konold, 1989). Students' ideas often suffer from misconceptions or misinterpretations of scientific models. Even students after a statistics course and trained statisticians may retain and use invalid intuitions (Díaz et al., 2010).

Modern probability education focuses on concepts that

> played a key role in the history and form the base for modern theory of probability while at the same time people frequently hold incorrect intuitions about their meaning or their application in absence of instruction. (Batanero, 2014, p. 493)

Such concepts include the ideas of random experiment and sample space, addition and multiplication rules, independence and conditional probability, random variable and distribution, combinations and permutations, convergence, sampling and simulation (Batanero, 2014). Similarly, Gal (2005) includes five big ideas as the building blocks of probability literacy: variation, randomness, independence, predictability or uncertainty. Of the important ideas mentioned above, the research presented in this paper will focus on two closely related: independence and conditional probability. Research has shown that even trained professionals sometimes poorly judge independence and conditional probability. For example, they understand conditioning as a causal relationship (Batanero & Sanchéz, 2005).

Misconceptions and misinterpretations in the perception of the independence of random events are well documented. For example, the gambler's bias and the hot hand fallacy (Roney, 2016) are grounded in a student's belief in the internal connection between consecutive events. We often meet fallacies related to independence when looking for all possible outcomes. Students make errors of order or repetition, provide a non-systematic listing of the sample space or make faulty interpretations of diagrams (Jones et al., 2007). Misconceptions in conditional probability are manifested in confusing independence and causality or the fallacy of the time axis, exchanging the role of events, reversing conditions and conditional statements,


https://doi.org/10.14712/18047106.3013
Nakladatelství Karolinum

confusing mutually exclusive events and independent events, etc. (Diaz et al., 2010; Batanero & Borovcnik, 2016).

This study is a part of qualitative research on students' comprehension and conceptualisations of probability and statistics. Specifically, the study focuses on gaining insights into understanding independence and conditional probability among Czech secondary and university students.

## 2 Theoretical framework and literature review

A brief historical and epistemological overview of probability will be given first to provide a wider context for students' interpretations. We will focus on two key concepts (independence and conditional probability) that are essential to learning probability and are our study's focus. Next, we will provide a literature review of studies focusing on students' conceptions of independence and conditional probability.

### 2.1 Conceptions of probability

The mathematical description of probability was provided in 1933 by Kolmogorov's axioms, but there are still extensive discussions about its meaning. Probability interpretation is far from resolved (Galavotti, 2017). Two perspectives on the nature and interpretation of probability can be distinguished – epistemological (with subjective and logical interpretations) and ontological (with frequency-based and propensity interpretations).

The epistemological conception of probability describes our relationship to reality. Randomness is only a lack of information, and probability is a tool for quantifying a lack of knowledge. There are two interpretations. The subjective interpretation views probability "as a personal degree of belief" (Batanero, 2015, p. 36) that a person assigns to an event and is based on their willingness to place a bet regarding the event. According to the logical interpretation, the degree of belief is generally accepted by all observers based on logical arguments. Its basic principle is the principle of indifference which assigns equal probability to all elementary events if there is no reason to prefer any of them.

In the ontological conception, randomness is a substantial part of physical reality and is independent of an individual's beliefs or subjective judgment. Part of the ontological perspective is frequency-based probability. It involves looking for the frequency of occurrence of an event when it is repeated many times. The statistical aspect of probability focuses on "objective mathematical rules through data and experiments" (Batanero, 2015, p. 36). The second branch consists of propensity interpretations based on the inclination or tendency of a situation to end with a given outcome.

### 2.2 Independence and conditional probability

This section will define independence and conditional probability.[1]

*Conditional probability* is a measure of the probability of a random event $A$ occurring, given that an event $B$ has already occurred: $P(A|B) \overset{\text{def}}{=} \dfrac{P(A \cap B)}{P(B)}$, $P(B) \neq 0$. Bayes' theorem applies: $P(B|A) = \dfrac{P(A|B) \cdot P(B)}{P(A)}$, $P(A) \neq 0$.

One way to define *independence* is by the relation in which the intersection is converted into a product. We define two random events $A$ and $B$ to be independent if and only if $P(A \cap B) = P(A) \cdot P(B)$. This definition is symmetric, and there is no need to assume a non-zero probability. The second definition is based on conditional probability. We define two random events $A$ and $B$, where $P(B) \neq 0$, to be independent if and only if $P(A) = P(A|B)$. Both definitions are interrelated, as can be seen from the equality $P(A \cap B) = P(A|B) \cdot P(B) = P(A) \cdot P(B)$.

From the educational point of view, the second definition is more illustrative as it points to the essence of independence. The independence of random events $A$ and $B$ means that information about one of them, e.g., $B$, does not affect the probability of event $A$. The probability of event $A$ does not depend on whether event $B$ has occurred.

[1] Independence was first dealt by founders of probability, such as Abraham de Moivre: "Two events are independent, when they have no connection one with the other, and that the happening of one neither forwards nor obstructs the happening of the other." (de Moivre, 1756, p. 6) Independence and conditional probability was thoroughly analysed by Thomas Bayes and completed by Pierre-Simon Laplace (Dale, 1982).

## 2.3 Students' perception of independence

The perception of independence manifests well in the expectation of results in a sequence of the same partial random trial. Understanding of independence is closely related to examples of repeated random events, such as coin flipping or dice tossing, spinning roulette, in the lottery, but also in the birth of a boy or a girl, in anticipation of a fall or rise in the prices of commodities on the stock exchange, etc. Such examples are often reflected in students' erroneous intuitive ideas about independence or conditional probability. Intuition leads them to abandon the idea of independence and use the pattern of past data to predict the next outcome.

Two well-documented misconceptions in this area are the gambler's fallacy and the hot hand fallacy (Roney, 2016), in which the law of large numbers plays a crucial role. It posits that the ratio of the outcomes corresponding to event $A$ and all the outcomes approaches the probability of event $A$. In the long-term repetition of trials, the probability of event $A$ can be approximately replaced by the actual ratio of the outcomes corresponding to $A$ and all the outcomes. However, this does not apply to a few repetitions. The tendency to make false conclusions about the probability based on a few realisations of the trial is called Kahneman's law of small numbers (Tversky & Kahneman, 1971).

While Piaget and Inhelder (1951) supposed that 15-year-old pupils understood the law of large numbers, later studies brought differing results and showed that people of all ages are prone to erroneous decisions even if they learned formal probability theory (Batanero, 2015). For example, the law of small numbers often influences the expectation of repeated trial outcomes. It manifests itself when the probability of an event is also applied to small numbers of its repetitions, regardless of the low reliability of such a conclusion. This mistake is deeply ingrained in gamblers and is often called the gambler's fallacy (bias). It is based on the mistaken belief that if a random event has occurred more frequently than usual, it will occur less frequently in the future. Similarly, the hot hand fallacy (Roney, 2016) is the belief that there is a greater chance of success after a series of successful trials. Graham explains possible reasons for this misconception:

> People sometimes appeal to the "law of averages" to justify their faith in the gambler's fallacy. They may reason that, since all outcomes are equally likely, in the long run, they will come out roughly equal in frequency. However, the next throw is very much in the short run and the coin, dice or roulette wheel has no memory of what went before. (Graham, 2006, p. 58)

The gambler's fallacy is usually explained by misapplications of heuristics regarding random sequences of events. People sometimes behave as if independent events were related. They are internally convinced that long runs of one outcome are unrepresentative and should be unlikely to occur (Tversky & Kahneman, 1971). Other perspectives on the gambler's fallacy emphasise its psychological aspects. For example, a long run of the same results motivates the participants to make certain conclusions and judgments (Roney, 2016). A long run of red numbers in roulette leads to the belief that it must be interrupted at some time. On the other hand, a long run of basket-shooting successes may increase the expectation of further successes with the justification that the player is simply doing well.

## 2.4 Students' perception of conditional probability

Two conceptions of conditional probability can be distinguished. If there is no causal relationship between events $A$ and $B$, we speak about the epistemological understanding of conditional probability. If a causal relationship exists between events $A$ and $B$, conditional probability is understood substantially: Event $A$ results from event $B$. We call such interpretations *ontological* and *causal*. If we perceive conditional probability causally, a problem arises as the direction of inference from cause to effect cannot be reversed. The casual perception of probability can lead to the so-called Humphreys Paradox, which demonstrates that conditional probabilities are formally symmetric, but cause and effect are not symmetrical (Mccurdy, 1996; Humphreys, 1985).

Mixing the conditional probability $P(A|B)$ and the causal relationship between events $A$ and $B$ is common even for university students (Batanero & Sanchéz, 2005; Díaz & de la Fuente, 2007; Díaz et al., 2010). Another frequently identified misconception is a student's expectation that event B should always precede event A (Batanero & Sanchéz, 2005), also called the fallacy of the time axis (Falk, 1989). Referring to international studies, Díaz et al. (2010) identified other misconceptions concerning independence and conditional probability, such as exchanging the events in conditional probability, confusing mutually exclusive events with independent events, etc.

Tarr and Lannin (2005) present an overview of research documenting the above misconceptions of students of different ages, including undergraduates. They distinguished four levels of students' thinking.[2]

[2]Similarly, Jones et al. (2007) propose four levels in the understanding of probability: subjective, transitional, informal quantitative, and numerical. At the highest level, for example, a student can distinguish independent and dependent events.

Level 1 is characterised by a subjective assessment of the situation according to students' beliefs and experiences. Such students can reason about certain or impossible events but ignore numerical data about probability. They consider consecutive events related and have no images of independence and conditional probability. Students with Level 2 thinking recognise that probability changes in some situations but do not assign numerical values to conditional probabilities. They can distinguish which events are related and which are not. Students with Level 3 thinking discern when one event affects another and can distinguish independence between them. They can quantify probability; however, they have shortcomings when considering independence and determining probability. Students with Level 4 thinking can assign numerical probability values to situations and know the conditions under which these hold. They can evaluate the probabilities of consecutive events and distinguish between dependent and independent events.

Mooney et al. (2014) present a synthesised framework of probabilistic thinking. The lowest is the level of prestructural probabilistic thinking. Students intuitively understand randomness and believe that consecutive events are always related. Their thinking is irrelevant, non-mathematical, or personalised. On the level of unistructural probabilistic thinking, students tend to revert to subjective probabilistic thinking, but they already compare probabilities or determine conditional probabilities. On the level of multistructural probabilistic thinking, student thinking is quantitative and proportional. They, for example, recognise changes in probability and independence in without-replacement events and use ratios, counts, probabilities or odds in judging probabilistic situations. Finally, on the level of relational probabilistic thinking, student thinking shows an interconnection of probabilistic ideas. Students, for example, determine probabilities for complex situations, including non-equally likely situations.

### 2.5 Local context and research questions

In the Czech Republic, probability is not part of the Framework Education Programme for Basic Education, which is the main curricular document, abiding for all schools. Probability only comes at the upper secondary school. According to the Framework Education Programme for Secondary General Education (Grammar schools), the pupils are expected to solve real problems with a combinatorial structure, utilise combinatorial analysis methods when calculating probabilities, discuss and critically evaluate statistical information, select and employ appropriate statistical methods and represent data sets graphically. The Framework is not very specific about the concepts which should be targeted: "random events and their probability, probability of the union and intersection of events, independence of events" (FEP, 2007, p. 24). Lessons on probability are usually taught in the final grades in Czech secondary schools. Before that, students can only be expected to have intuitive ideas about independence and conditional probability. Both concepts are present in Czech mathematics textbooks for secondary schools (Mošna, 2022). Independence is usually defined with the help of the definition of conditional probability. Some textbooks also introduce the independence of three events.

Courses on statistics and probability are common at universities, focusing on applying concepts at technical schools or understanding big ideas at the faculties educating mathematics teachers (Mošna, 2022). The international research introduced in the sections above has shown that many misconceptions and intuitive ideas remain even after a formal education on probability. The question arises whether the same applies to Czech students.

Drawing on the studies above, the present study focuses on the following questions:

1. How do secondary and university students perceive independence in tasks requiring multiple repetitions of random events? To what extent are known misconceptions manifested in their considerations?

2. How do secondary and university students perceive conditional probability? Are there signs of epistemological and causal understanding in their considerations?

## 3 Methodology

### 3.1 Sample and research design

The respondents of the study are Czech secondary school and university students. The sample is convenient. Secondary school and university students were invited to participate in the study by the author. They were the students of the universities where he worked (Czech University of Life Sciences in Prague, Faculty of Education of Charles University in Prague) and graduates and students of a Prague secondary school. The sample was complemented by students from the author's social circle. A heterogeneous sample of students who had not undergone the formal teaching of probability and those who had completed such teaching was sought.

Complete anonymity, protection of all personal data and disposal of all documents after the research was guaranteed to the participants. They were informed that the purpose of their solution of tasks was research, and their answers could not adversely affect them. The names mentioned below are fictitious pseudonyms.

A total of 43 participants (18 females, 25 males) participated in the study. They were divided into three groups:

- Group I – secondary school students with no teaching of probability and statistics (aged 14–18), $N = 12$.
- Group II – secondary school or university students who underwent teaching probability and statistics at secondary schools but had not met these topics at the university (aged 18–20), $N = 17$.
- Group III – university students who had taken a course on probability and statistics (aged over 20), $N = 14$.

Given the research aim and questions, a qualitative research paradigm was used. As the aim was to get an insight into students' images of independence and conditional probability while solving tasks, individual task-based interviews were used. It was considered appropriate as such interviews are "intended to elicit in subjects estimates of their existing knowledge, growth in knowledge, and also their representations of particular mathematical ideas, structures, and ways of reasoning" (Maher & Sigley, 2014, p. 579).

The author conducted the interviews. First, the students were asked: "What do you think the independence of two random events means?" Next, they solved tasks on independence and conditional probability and were asked to think aloud while solving them. One by one, the students were presented with the tasks and had time to think about their solutions. The interviewer did not interfere unless the student was stuck. He provided a hint in such a case. He also asked clarifying questions if the student's statement was not clear. The interviews lasted 30 to 45 minutes. They were video-recorded and transcribed to be analysed.

Carefully constructed tasks are key components of the task-based interview (Maher & Sigley, 2014). The author prepared the tasks focusing on students' understanding of independence and conditional probability based on research literature (Batanero & Sanchéz, 2005; Díaz & Batanero, 2009; Konold et al., 1993). The tasks were piloted with 14 students outside the sample to ensure the comprehensibility and unambiguity of their formulations.

## 3.2 Research tool

The research tool consisted of eight tasks. The first six tasks focused on the perception of independence, and the last two on the perception of conditional probability. The coins and dice were supposed to be symmetrical in the tasks. It was presumed that students had an intuitive understanding that with repeated coin tosses, the result of each toss cannot be predicted, heads and tails have the same chance, heads and tails will occur approximately equally often, and the results of heads and tails alternate unsystematically and irregularly.

Next, we present an *a priori* analysis of all the tasks.

> Task 1: Peter flips two ten-crown coins. What is the probability that both coins land on heads? Dusan flips two coins – a ten-crown and a five-crown. What is the probability that both coins land on heads?

The student must decide whether the set of possible outcomes consists of ordered pairs of heads and tails (permutations) or unordered ones (combinations). The toss of two different coins can be divided into two independent tosses of a ten-crown coin and a five-crown one. The probability that the ten-crown coin lands on heads is $P(A_H) = 1/2$, and similarly for the five-crown coin: $P(B_H) = 1/2$. The probability that both coins land on heads is $P(A_H \cap B_H) = P(A_H) \cdot P(B_H) = \dfrac{1}{4}$. A tree diagram in Figure 1 depicts the situation. The student must realise that the coins behave the same regardless of what is written on them; in other words, they are always distinguishable. That is why Peter and Dusan get the same results.

Tasks 2 and 3 generalise Task 1 to the twofold multiple-outcome trial and the fivefold two-outcome trial. They aim to determine how much task complexity affects the understanding of independence.

> Task 2: We roll two dice. Is the probability of the outcome of five and six equal to the probability of the outcome of two sixes?

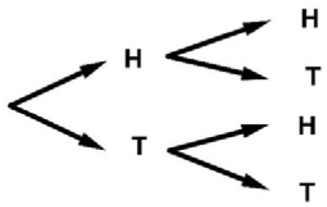



**Fig. 1:** Tree diagram for Task 1

Similarly to Task 1, the set of all outcomes is formed by all ordered pairs of numbers 1 to 6 with repetition. Therefore, the probability of rolling five and six is 1/18 (pairs $[5, 6]$ and $[6, 5]$), and the probability of rolling two sixes is 1/36 (pair $[6, 6]$). The student might intuitively suppose that the probabilities of six and five and two sixes are equal.

Task 3: We toss a coin five times. Which of the following outcomes has the lowest probability? Outcomes: a) head, tail, head, tail, head, b) head, head, head, head, head, c) head, tail, tail, head, tail.

A tree diagram shows that the result of an individual toss is not affected by the previous result. Thus, the probability of all three possibilities is the same, namely 1/32. Answer a) may indicate a student's false belief in the irregular alternation of heads and tails, while answer b) may indicate the gambler's bias.

Task 4: We toss a coin five times. Which of the following results has the greatest probability? We will have: a) heads on three coins, tails on the other two coins, b) heads on all five coins, c) heads on two coins, and tails on the other three coins.

Task 4 is similar to Task 3, but the order of heads and tails does not matter this time. The result in case (b) is 1/32 and in (a) and (c) 5/16. We can use a tree diagram again. To use the solution based on a set of all outcomes, it is necessary to consider an ordered five-tuples of heads and tails as a single outcome.

Successful solutions to Tasks 3 and 4 demonstrate a student's ability to distinguish the specific order of individual outcomes from the total number of outcomes with the same portion of heads and tails.

Tasks 5 and 6 focus on conditional probability and independence. The inductive method (from the specific to the general) is used in Task 5, and the deductive method (from the general to the specific) is used in Task 6. Task 5 falls within the classical interpretation of probability, while Task 6 calls for subjective interpretation.

Task 5: We toss a coin repeatedly. The coin has landed on heads ten times in a row. What is the chance that the eleventh outcome is heads again? a) Very small. b) Very large. c) Greater than 1/2. d) Equal to 1/2. e) Less than 1/2.

Again, as the independence of individual tosses is assumed, the outcome in the eleventh trial is not affected by any of the first ten outcomes. The probability of the outcome "heads in the eleventh tossing" is the same as in any other trial (1/2). It is also important to realise that we use the principle of indifference in the frame of classical (or logical) interpretation of probability. The symmetry of the coin is given.

Task 6: Peter and Dusan play table tennis and Dusan has won ten times in a row. What is the chance that Dusan will also win in the eleventh match? a) Very small. b) Very large. c) Greater than 1/2. d) Equal to 1/2. e) Less than 1/2.

The probabilities of Peter's or Dusan's win need not be equal, and they can even change. We can estimate the next result based on our previous experience. Dusan can probably play table tennis much better than Peter (option b) or c)). Moreover, Dusan's and Peter's performance can vary in each game; they can learn from their mistakes and use the opponent's weak sides in the next match.

Task 7: Out of a box with four balls, two black and two white, Dusan draws one ball and puts it in his pocket, and then Peter similarly draws one ball and puts it in his pocket. What is the probability that Dusan has a black ball in his pocket? What is the probability that Peter has a black ball in his pocket? How does the probability that Petr has a black ball in his pocket change when Dusan shows that he has a black ball in his pocket? What is the probability that Dusan had a black ball in his pocket when Peter drew a black one?

The answers to the first two questions are the same (1/2), even though Dusan draws first. Students may not realise that and may say they cannot answer because they do not know what Dusan has drawn. They might conclude that the probability is 1/3 if Dusan has drawn a black ball and 2/3 if Dusan has drawn a white ball. Such an answer would suggest a causal understanding of probability. If, on the other hand, the student accepts an epistemological explanation, they can use the total probability theorem[3] or a tree diagram to solve the task.

The answer to the third question is easy (1/3). Students might be surprised that the fourth question has the same result. If they understand conditional probability causally, they will conclude that Dusan drew the black ball with probability 1/2, regardless of Peter's following outcome. If they understand it epistemologically, they can determine the correct probability of 1/3.

> Task 8: There are two ordinary dice in the box (numbered 1, 2, 3, 4, 5, 6) and one special dice where five is replaced by six (numbered 1, 2, 3, 4, 6, 6). Dusan draws one of these dice at random. He throws it. What is the probability that a six will be thrown? Peter does not see the whole process. He thinks about what kind of dice Dusan has probably thrown. What is the probability that Dusan has thrown a special type of dice?
> Dusan reports rolling a six. What is the probability (for Peter) that Dusan has rolled a special dice? How would the situation change if Dusan reported that he rolled a five?

This task is similar to Task 7. The total probability theorem yields the probability that Dusan rolls a six: $P(A) = 1/6 \cdot 2/3 + 2/6 \cdot 1/3 = 2/9$. Some authors suggest performing Bayesian calculations with absolute frequencies (Martignon & Wassner, 2002): there are four sixes out of 18 sides (the total number on all three dice), thus $P(A) = 2/9$. The probability that Dusan drew the special dice is $P(B_2) = 1/3$. To calculate how this probability changes after additional information that Dusan has rolled a six, we can use Bayes' formula (see Section 2.2) $P(B_2|A) = \frac{\frac{1}{3}\cdot\frac{1}{3}}{\frac{2}{9}} = \frac{1}{2}$ or the definition $P(B_2|A) = \frac{P(A \cap B_2)}{P(A)}$, $P(B_2|A) = \frac{\frac{2}{18}}{\frac{2}{9}} = \frac{1}{2}$. A simple consideration also leads to a solution: two of the four possible sixes are on the special dice, and two are on the ordinary dice (one six on each of the two ordinary dice).

The last question of Task 8 can uncover the epistemological nature of probability. The probability that Dusan had drawn the special dice when he rolled a five can be found by simple reasoning. This probability is 0 because five is not on the special dice.

## 3.3 Data analysis

The analysis of the interview transcripts was supplemented by two additional sources of data: the students' artefacts, specifically their solutions to the tasks, and the field notes taken by the interviewer (author) during the interviews.

The accuracy of the students' answers for each task and individual was tracked during the initial analysis. Subsequently, a second round of analysis was conducted, wherein the data were coded using a preliminary coding framework developed based on an a priori analysis of the tasks in the research tool. Through iterative data readings, new codes emerged. This process led to establishing a coding framework comprising codes and their corresponding descriptions, drawing on the methodology outlined by Saldaña (2015). The codes encompassed the anticipated misconceptions, incorrect reasoning, and expected justifications for the conclusions drawn by the students. In the subsequent stage, the codes were further categorised into two themes, independence and conditional probability, aligning with the research questions posed in the study.

**Independence:**

Correct solution: Graphs, Sample spaces, Variations, Combinations
Incorrect solution: Faulty interpretation of graph, Faulty sample spaces (Errors of repetition, Errors of order), Incorrect use of variations, Incorrect use of combinations
Misunderstanding the question

Conditional probability:

Correct solution: Conditional probability, Total number of cases
Incorrect solution: Exchanging the events, Fallacy of the time axis, Non-adequate notation
Misunderstanding the question
Interpretation of probability: Rather epistemic, Rather casual

[3] Let us suppose that the set of all outcomes ? can be factorized into two parts $B_1$ and $B_2$ which cover it ($B_1 \cup B_2 = \Omega$), which are disjoint ($B_1 \cap B_2 = \emptyset$), and moreover $P(B_1) \neq 0$, $P(B_2) \neq 0$. Then for random event $A \subset \Omega$, $P(A) \neq 0$, it holds $P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2)$.

For example, the statement in which the student would not answer the question of which dice was originally drawn if a six subsequently fell in Task 8 was coded '**Conditional probability**: Casual understanding, Fallacy of the time axis'.

The author coded the data in two rounds, and the coding consistency was checked by a collaborator in about 20% of the data. Next, the frequency of codes was calculated for Group I, Group II and Group III.

# 4 Results

The results will be presented for each research question separately. They will be illustrated by the students' quotes from the interviews.

## 4.1 Perception of independence

First, we will look at the number of correct answers in each group (Tab. 1). There seems to be a slight improvement in understanding independence with the level of study. However, the sample is small, and we cannot make any conclusions about differences between groups regarding the success rate. Thus, we will present the results for the whole sample.

**Tab. 1:** Absolute and relative frequencies of correct answers for Tasks 1–6 per group

| | Group I ($N = 12$) | | Group II ($N = 17$) | | Group III ($N = 14$) | | In total ($N = 43$) | |
|---|---|---|---|---|---|---|---|---|
| Task 1 | 8 | 66.7% | 15 | 88.2% | 14 | 100.0% | 37 | 86.0% |
| Task 2 | 3 | 25.0% | 5 | 29.4% | 4 | 28.6% | 12 | 27.9% |
| Task 3 | 1 | 8.3% | 4 | 23.5% | 4 | 28.6% | 9 | 20.9% |
| Task 4 | 2 | 16.7% | 6 | 35.3% | 9 | 64.3% | 17 | 39.5% |
| Task 5 | 4 | 33.3% | 12 | 70.6% | 12 | 85.7% | 28 | 65.1% |
| Task 6 | 7 | 58.3% | 16 | 94.1% | 14 | 100.0% | 37 | 86.0% |

In the interviews, students characterised independence as a relationship when two events do not influence each other. They said, for example, that independence "is a state in which the subject is completely unaffected by another", "is when the elements do not interact with each other", or "is the absence of a relationship between two variables". Some students provided a general answer, saying that independence is "freedom", "self-sufficiency", or "non-dependence".

Next, we will present results related to the tasks used in our study.

In Task 1, we focused on the agreement or difference between two situations: tossing two identical and two different coins. Most respondents (37 out of 43) considered it the same. Their arguments were "both cases are identical, coins as coins" or "the value of the coins does not matter". However, only 32 of them could calculate the probability correctly. Jirka's approach (Group I) is worth attention. He initially concluded: "The probability is 1/4 if we toss ten-crown and five-crown coins, and the probability is 1/3 if both coins are ten-crown ones." Then, without the interviewer's prompt, he changed his mind and added to himself: "But no two coins are identical." After this the interviewer initiated a conversation about how the coins could be different. He told Jirka that coins could differ in colour and attempted to lead him to a paradoxical situation where coins would behave differently for a colour-blind observer and a person with normal colour vision. In the end, Jirka insisted that "it depends on the assignment; if the coins are different, the correct probability is 1/4, and if we guarantee they are not different, then 1/3". This indicates Level 1 in Tarr and Lannin's (2005) model.

Mathematically, the solution of Task 2 is similar to that of Task 1, yet many students did not get the correct result. Less than a third of them (12 out of 43) realised that the probabilities of rolling a six and a five differ from the probability of rolling two sixes. In addition, some students gave the correct answer (that the probabilities are different) but based on misconceptions, e.g., that "five rolls more often". A slight complication of the task thus brought completely incorrect considerations.

Task 3 is a generalisation of Task 1 from the mathematical point of view, yet, its solution presented a problem for most students. The correct answer (the probability is the same in all cases) was provided by a few of them (9 out of 43). Outcome b) was considered the least probable, which was expected.

In Task 4, many students realised the connection with the previous task. They found that the total number of tails and heads differed from only one individual outcome (with the same ratio). Alena (Group III) formulated it well: "It looks the same as the last task, but it is not. There seems to be an omission of arrangement, which is a very important aspect of probability." The interviewer confirmed her answer. There were more correct answers than in the previous task, but not by many. Few respondents (17 out of 43) realised that the probability of results a) and c) is greater than the probability in b). Only nine respondents considered a) the most probable, and 12 selected c). No student opted for b). Five out of 43 respondents considered all three results equally probable.

Task 5 asks for the probability that after ten heads, the next outcome will be heads again. 28 students provided the correct variant d). However, it also means that more than a third of students in the sample do not perceive the independence of individual outcomes. Alena (Group III) justified her decision for a) 'Very small': "[The probability] is small, but the chance is a jerk." Another student supported his option e) 'Less than 1/2' by commenting: "…it would be very regular."

Task 6 describes Dusan's chance of winning in the eleventh game after winning the ten previous ones. Most students realised that Dusan is probably a better player and chose b) 'A high probability' (29 out of 43 respondents) or c) 'Greater than 1/2' (8 out of 43 respondents). Contrary to Task 5, a subjective interpretation of probability is needed. An example of students' reasoning is: "It's about [Dusan's] abilities." The subjective conception was also reflected in the statements "[Am I] a bookmaker?" or "I support Peter". Only 6 out of 43 students used the principle of independence and indifference and answered d) 'Equal to 1/2'. Their answers correspond to a subjective perception of probability based on the individual's willingness to bet and may be different for everyone. Peter and Dusan's matches do not have to be independent. They are interconnected and influenced by both players' form and game development.

## 4.2 Perception of conditional probability

The last two tasks focused on students' perception of conditional probability. Table 2 shows that indications of epistemological and causal interpretations appeared in all the groups and that, for many students, it was impossible to decide on the type of interpretation.

**Tab. 2:** Numbers of interpretations of conditional probability in Tasks 7 and 8

| | Group I | Group II | Group III |
|---|---|---|---|
| Indications of epistemological interpretation | 1 | 4 | 5 |
| Indications of casual interpretation | 3 | 8 | 6 |
| It was not possible to decide | 8 | 5 | 3 |

The students' answers to Task 7 fully confirmed the expectation of the causal interpretation. Almost all answered the first question (what is the probability that Dusan will draw a black ball) correctly (40 out of 43). The second question (what is the probability that Peter will draw a black ball) was only answered by 30 students. They missed information about what Dusan had drawn. Alena (Group III) expressed her helplessness: "Am I a clairvoyant?" Some students reached the results of 1/3 when Dusan drew a black ball and 2/3 when he drew a white ball; however, they did not take into account the overall probability of 1/2. Such students answered 1/3 to the third question (how the probability that Peter drew a black ball changes when we know that Dusan drew a black ball). In total, 23 students correctly answered the overall probability of 1/2.

On the contrary, 7 students claimed that the probability that Peter would draw a black ball was the same, whether we know what Dusan had drawn. These students probably understand probability causally and not epistemologically.

The difference between the epistemological and causal conceptions was manifested in the fourth question of Task 7. The relationship between what Dusan drew and what Petr drew can be understood causally or epistemologically. Alena (Group III) stated: "I am slowly losing myself. If Dusan drew first, what Peter drew is irrelevant, right?" She did not realise the relationship between the two events, which indicates Level 2 in Tarr and Lannin's (2005) model. Less than a third (14) of students gave the correct answer.

Task 8 was difficult for the study sample, as only five students answered all the questions correctly. Others often said that the task was beyond their ability and imagination: "I'm at my wits' end", "I don't know anymore", and the like. Most students perceived the epistemological nature of probability as problematic: "The situation is different only psychologically; the probability is the same." (Pavel, Group III)

# 5 Discussion

The study confirmed students' problems with understanding the concept of independence as known from the literature (Díaz et al., 2010; Roney, 2016). In the tasks on independence (1–6), to an extent, the students could create sample spaces for two-stage experiments, which indicates that their probabilistic thinking is on the unistructural level (Mooney et al., 2014). However, their arguments were often incorrect, and they had great difficulties if the tasks got slightly more complicated. Thus, the students cannot construct such sample spaces systematically and have not reached higher levels of probabilistic thinking (multistructural or relational, ibid.).

Students' somewhat vague ideas about independence mainly manifested themselves in tasks related to repeating partial random events. Incorrect conclusions included assigning a higher probability to randomly distributed irregular $k$-tuples of outcomes and a lower probability to $k$-tuples where partial events were repeated or were periodic. The students often missed the principle that, in the case of independence, the probability of one event is not affected by the results of the previous (or subsequent) event.

The study focused on how students' lack of understanding of independence affects the method of estimation and probability calculations in specific cases with repeating trials. The understanding of independence and multiple repetitions of a random event has been addressed by several studies (Batanero & Sanchéz, 2005; Konold et al., 1993). They conclude that students can successfully solve simple tasks, such as repeating a random experiment twice with two outcomes (a double toss of a coin or a toss of two coins). The study presented in the paper showed that problems mainly occurred in more complicated tasks. The students realised that the outcomes of two simple trials were independent and correctly calculated their probability. However, if the task got slightly more complicated, misconceptions about independence occurred, and students could not calculate the correct probabilities. For example, a double experiment of an event with multiple outcomes (rolling two dice) or multiple repetitions of a simple event (tossing five coins) were shown to be problematic tasks for students.

We also showed that the incorrect understanding of the repeated results is often related to the incorrect understanding of the assignment. There is an important difference between a certain sequence (or 5-tuple) of heads and tails with their order on the one hand and a group of outcomes with a certain total number of heads and tails on the other hand. The discrepancy between intuitive perception and calculations seems to be related to the formulation of the question and the answer we are looking for. For example, Task 5 asks for the probability of tossing heads in the eleventh trial, i.e., the probability of tossing heads in one trial after tossing heads in the previous ten rolls. It means something other than the probability that heads will occur in all eleven tosses.[4] Many students found it impossible to accept that the probability of heads is still 1/2 after ten heads. The intuition of randomness is more suited to a certain irregularity and rotation of partial outcomes with an approximate ratio of $1:1$. On the contrary, in a similar case (in Task 6), the students realised that the probability that Dusan would win again after ten previous wins was relatively high because we judge it based on a subjective conception of probability and use an inductive way of reasoning.

In the conditional probability tasks (7, 8), the students could calculate the conditional probability of the events that followed the condition. It indicates an unistructural level of probabilistic thinking (Mooney et al., 2014). Nevertheless, they could not calculate the conditional probabilities when the event had pre-conditions. They could not recognise this situation change, and, thus, their probabilistic thinking remained on the unistructural level.

Several reasons for the incorrect perception of conditional probability are usually discussed in the literature. First, it is a causal perception of conditional probability, its temporal perception, the confusion of considered events and other misconceptions (Díaz & de la Fuente, 2007; Díaz & Batanero, 2009; Tanujaya et al., 2018). In this study sample, conditional probability was mostly perceived as causal. However, epistemological ideas are not related to random events but to information about them. Information about what happened before can be influenced by information about what followed. A two-way epistemological relationship can replace the one-way causal relationship in these considerations. This may lead to some inconsistencies. Cause and effect cannot be confused. The formula for calculating conditional probability can be applied to both sides, i.e., for $P(A|B)$ and for $P(B|A)$ but must be understood epistemologically. In the presented study, the perception of conditional probability based on the position on the timeline was evident. This phenomenon has a notable influence on the causal perception of conditional probability.

Despite the limited sample size and the relatively small number of students within each group, the findings suggest that misconceptions were present in the solutions provided by students across different groups, regardless of their prior exposure to probability education. These findings align with international research, highlighting the propensity for misunderstanding and reliance on intuitive judgments among even experienced professionals (Batanero & Sanchéz, 2005; Díaz et al., 2010).

However, it is important to acknowledge the limitations of our study. Firstly, the small sample size necessitates caution in generalising the results. They should be interpreted in the context of a restricted sample. While we cannot make broad claims, the study does offer insights into how students approach the concepts of independence and conditional probability and how their intuition can lead to misconceptions. A larger sample would be required to validate the uncovered phenomena.

[4] In the former case, it is 1/2, and in the latter 1/2048. The former consists of the situation with one toss, and the latter concerns the situation with 11 tosses. Limit theorems can be applied in the latter case but not in the former.

The selection of tasks employed in this study introduces another limitation. Different tasks might provide a more nuanced understanding of students' reasoning in this domain. Further research examining students' perceptions of independence, conditional probability, and probability in general would be beneficial in expanding the knowledge in this area.

## 6 Conclusions and implications

The findings of this study confirmed that students, including those who had undergone a university course on probability, encountered difficulties in comprehending the concepts of independence and conditional probability. Notably, misconceptions regarding independence became apparent in more challenging tasks, wherein students relied more on their intuition than their acquired knowledge. Specifically, students exhibited confusion between scenarios involving the total number of events and those involving a specific sequence of outcomes, including their order. The misconceptions in perceiving independence were particularly evident when describing the random space, encompassing the complete set of outcomes in a random event.

The study has some educational implications. Firstly, it is advisable to solidify students' foundational understanding of independence through ample elementary examples, gradually progressing to tasks of increasing difficulty. When a student encounters difficulty with a complex problem, it can be beneficial to compare the results obtained in easier tasks and those in more challenging ones. Secondly, students should be encouraged to analyse the problem at hand from different angles. For instance, when considering the tossing of five coins, highlighting the distinction between two types of questions can be helpful: one concerning the probability of a specific sequence (or quintuple) of heads and tails in a given order, and another about the probability of a group of outcomes with a predetermined total count of heads and tails. Thirdly, the principles for solving problems in classical probability rely on symmetry. One can extend these principles by utilising symmetry and independence to encompass scenarios involving double or multiple throws or moves. It is often advantageous to decompose random experiments into individual steps. Once the correct calculation methodology becomes sufficiently clear, concepts such as permutations or combinations and the application of combinatorial formulas can be introduced. Experimental verification can further reinforce the obtained results. Lastly, graphical representations, such as tree diagrams, proved effective in delineating the decomposition above into simple partial steps as they illustrate fundamental probability concepts (Díaz et al., 2010; Graham, 2006; Rolka & Bulmer, 2005).

## References


Albert, J. H. (2003). College students' conceptions of probability. *The American Statistician*, *57*(1), 37–45. https://doi.org/10.1198/0003130031063

Batanero, C. (2014). Probability teaching and learning. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 491–496). Springer. https://doi.org/10.1007/978-94-007-4978-8_128

Batanero, C. (2015). Understanding randomness: Challenges for research and teaching. In K. Krainer, & N. Vondrová (Eds.), *Proceedings of the 9th Congress of European Society for Research in Mathematics Education* (pp. 34–49). Charles University in Prague, Faculty of Education and ERME.

Batanero, C., & Borovcnik, M. (2016). *Statistics and probability in high school.* Springer. https://link.springer.com/book/10.1007/978-94-6300-624-8

Batanero, C., & Sanchéz, E. (2005). What is the nature of high school students' conceptions and misconceptions about probability? In G. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 241–266). Springer. https://doi.org/10.1007/0-387-24530-8_11

Dale, A. I. (1982). Bayes or Laplace? An examination of the origin and early applications of Bayes' theorem. *Archive for History of Exact Sciences*, *27*(1), 23–47. https://doi.org/10.1007/BF00348352

de Moivre, A. (1756). *The doctrine of chances.* (3rd ed.). A. Millar.

Díaz, C., & Batanero, C. (2009). University students' knowledge and biases in conditional probability reasoning. *International Electronic Journal of Mathematics Education*, *4*(3), 131–162. https://doi.org/10.29333/iejme/234

Díaz, C., Batanero, C., & Contreras, J. M. (2010). Teaching independence and conditional probability. *Boletin de Estadistica e Investigacion Operativa, 26*(2), 149–162.

Díaz, C., & de la Fuente, E. I. (2007). Assessing students' difficulties with conditional probability and Bayesian reasoning. *Journal on Mathematics Education*, *2*(3), 128–148. https://doi.org/10.29333/iejme/180

Evans, B. (2007). Student attitudes, conceptions, and achievement in introductory undergraduate college statistics. *The Mathematics Educator*, *17*(2), 24–30.

Falk, R. (1989). Inference under uncertainty via conditional probabilities. In R. Morris (Ed.), *Studies in mathematics education: The teaching of statistics* (Vol. 7, pp. 175–184). UNESCO.

Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children.* D. Reidel

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic: Intuitively based misconceptions. *Journal for Research in Mathematics Education*, *28*(1), 96–105. https://doi.org/10.2307/749665

FEP (2007). *Framework Education Programme for Secondary General Education (Grammar Schools).* VUP.

Gal, I. (2005). Towards "probability literacy" for all citizens: Building blocks and instructional dilemmas. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 39–63). Springer. https://doi.org/10.1007/0-387-24530-8_3

Galavotti, M. C. (2017). The interpretation of probability: Still an open issue? *Philosophies*, *2*(3), 20. https://doi.org/10.3390/philosophies2030020

Graham, A. (2006). *Developing thinking in statistics.* SAGE Publications Inc.

Humphreys, P. (1985). Why propensities cannot be probabilities. *The Philosophical Review*, *94*(4), 557–570. https://doi.org/10.2307/2185246

Jones, G. A., Langrall, C. W., & Mooney, E. S. (2007). Research in probability: Responding to classroom realities. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 909–955), Charlotte.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction*, *6*(1), 59–98. https://doi.org/10.1207/s1532690xci0601_3

Konold, C., Pollatsek, A., Well, A., Lohmeier, J. H., & Lipson, A. (1993). Inconsistencies in students' reasoning about probability. *Journal for Research in Mathematics Education*, *24*(5), 393–414. https://doi.org/10.2307/749150

Maher, A. A., & Sigley, R. (2014). Task-based interviews in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 579–582). Springer. https://doi.org/10.1007/978-94-007-4978-8

Martignon, L., & Wassner, C. (2002). Teaching decision making and statistical thinking with natural frequencies. In B. Phillips (Ed.), *Proceedings of the 6th International Conference on Teaching of Statistics* [CD]. International Statistical Institute.

Mccurdy, Ch. S. I. (1996). Humphrey's paradox and the interpretation of inverse conditional propensities. *Synthese*, *108*(1), 105–125. https://doi.org/10.1007/BF00414007

Mooney, E. S., Langrall, C. W., & Hertel, J. T. (2014). A practitioner perspective on probabilistic thinking models and frameworks. In E. Chernoff, & B. Sriraman (Eds.), *Probabilistic thinking. Advances in mathematics education* (pp. 495–507). Springer. https://doi.org/10.1007/978-94-007-7155-0_27

Mošna, F. (2022). *Pojetí pravděpodobnosti a statistiky ve výuce.* [Conceptions of probability and statistics in teaching.] Pedagogická fakulta, Univerzita Karlova.

Nemirovsky, I., Giuliano, M., Pérez, S., Concari, S., Sacerdoti, A., & Alvarez, M. (2009). Students' conceptions about probability and accuracy. *The Montana Mathematics Enthusiast*, *6*(1–2), 41–46. https://doi.org/10.54870/1551-3440.1132

Piaget, J., & Inhelder, B. (1951). *La genése de I'idée de hasard chez I'enfant.* Presses Universitaires de France.

Rolka, K., & Bulmer, M. (2005). Picturing student beliefs in statistics. *Zentralblatt für Didaktik der Mathematik*, *37*(5), 412–417. https://doi.org/10.1007/s11858-005-0030-4

Roney, C. (2016). Independence of events, and errors in understanding it. *Palgrave Communications*, *2*, 16050. https://doi.org/10.1057/palcomms.2016.50

Saldaña, J. (2015). *The coding manual for qualitative researchers.* Sage.

Tanujaya, B., Prahmana, R. Ch. I., & Mumu, J. (2018). Designing learning activities on conditional probability. *Journal of Physics: Conference Series*, *1088*, 012087. https://doi.org/10.1088/1742-6596/1088/1/012087

Tarr, J. E., & Lannin, J. K. (2005). How can teachers build notions of conditional probability and independence? In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 215–238). Springer. https://doi.org/10.1007/0-387-24530-8_10

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Review*, *76*, 105–110. https://doi.org/10.1037/h0031322

# Pre-service teachers' noticing: On the way to expert target

Naďa Vondrová[1], Magdalena Novotná[1], Lenka Pavlasová[1], Jarmila Robová[2], Jana Stará[1], Klára Uličná[1]

[1] Faculty of Education, Charles University, M. Rettigové 4, 116 39 Praha 1, Czech Republic; nada.vondrova@pedf.cuni.cz
[2] Faculty of Mathematics and Physics, Charles university, Sokolovská 83, 186 75 Praha 8, Czech Republic

Unlike prevailing research focusing on what pre-service teachers attend to in a lesson and how they interpret it, the study investigates the content of their comments, knowledge-based reasoning and whether it agrees with experts' views. Study 1 determined the dimensions of quality teaching pertinent to lessons in which a new subject matter is introduced and made a noticing target. In Study 2, pre-service teachers ($N = 174$) at the end of their university study made a written reflection of a video lesson, which was compared against the target. Most could not discern situations important for deep work with the content in the lesson. They failed to apply their theoretical knowledge in interpreting the ones they mentioned. Only half of their comments included knowledge-based reasoning, and their views were mostly partially consistent or inconsistent with the experts' ones. This highlights the need to focus on content-related important situations in a lesson and their interpretation in teacher preparation and on developing the ability to discern the dimensions of instructional quality in concrete lessons.

Key words:
teacher noticing, knowledge-based reasoning, pre-service teachers.

Received 3/2023
Revised 5/2023
Accepted 6/2023

## 1 Introduction

The concept of professional vision capturing noticing and reasoning skills is a mediator between teachers' dispositions and classroom practice (Blömeke et al., 2015). The skills to notice effective teaching manifestations have been shown to correlate with implementing them in one's teaching, even for pre-service teachers (Sun & van Es, 2015; Wiens et al., 2021). This makes noticing an important target for teacher preparation, yet, pre-service teachers' (PST) noticing is lacking. PSTs tend to focus on management, pay little attention to content, and notice the teacher's rather than the pupils' actions. They tend to make general comments, evaluate rather than interpret and use naïve assumptions rather than theory for explanations even though they meet theory in their university courses (e.g., McDonald, 2016; Schäfer & Seidel, 2015; Simpson et al., 2018; Sonmez & Hakverdi-Can, 2012).

Studies on noticing use a normative frame of reference of what participants should notice to demonstrate noticing mostly implicitly. Only some studies present an explicit frame and investigate whether PSTs notice teaching-learning situations deemed important by experts and whether what they say is compatible with experts' views. We maintain that both aspects are relevant for PSTs' learning as future teachers. Thus, our research falls within the studies in which the instrument's validity is "grounded in the collective expertise of a community of experts" (Roose et al., 2018, p. 73). Taking expert noticing and knowledge-based reasoning as a long-term target, the paper aims to show how PSTs' noticing and knowledge-based reasoning at the end of their university study compares to that of experts.

## 2 Theoretical framework and literature review

The term 'professional vision' was coined as the expert ability to perceive and identify phenomena in a scene compared to the ability of lay persons (novices) (Goodwin, 1994). In education, it often overlaps teacher noticing, defined as "*professional vision* in which teachers selectively attend to events that take place and then draw on existing knowledge to interpret these noticed events" (Sherin et al., 2011, pp. 80–81).

While studies vary in their conception of professional vision, they agree on its two subprocesses (Blomberg et al., 2011; Sherin & van Es, 2009; Sherin et al., 2011): attending to events in an instructional setting (selective attention) and making sense of them (knowledge-based reasoning). The latter involves "processes of making sense of what has been noticed by linking observed situations to knowledge" (Schäfer & Seidel, 2015, p. 38) about teaching and learning. Later, this conception was expanded by acknowledging that to identify noteworthy features, one must disregard some features, too, and that to use one's knowledge to make sense of situations, one adopts a stance of inquiry (van Es & Sherin, 2021).

Analytical frameworks used to describe participants' reasoning differentiate whether the observer describes and/or evaluates the event or whether they also explain it (e.g., Sherin & van Es, 2009). We build on Stockero (2008), who distinguished describing, explaining, theorising, confronting (considering alternative explanations and others' points of view), and restructuring (re-examining one's beliefs and assumptions). The alteration, modifying a teaching-learning situation to reach its goal in a new/more effective way, is sometimes included in the frameworks capturing professional vision (Santagata, 2011). "Suggesting alterations within the [teaching-learning] situations is a way of professional learning" (Janík et al., 2019, p. 188) and is considered a sign of expertise (Stürmer et al., 2013).

Analytical frameworks do not usually differentiate between comments regarding the plausibility of the interpretation presented in these comments. For example, two comments may be coded 'theorise', as both include theory elements, but one may not be considered plausible by experts. Schäfer and Seidel (2015, p. 36) note:

> a teacher might notice an event and reason that student thinking was encouraged in the video [. . . ] but an expert in the field of teaching and learning viewing the same video would reason that student thinking actually was not being encouraged.

This study addresses the lack of research focus on the plausibility of PSTs' comments when reflecting on a lesson.

## 2.1 PSTs' noticing

Research brings converging results in that PSTs tend to focus on management and behaviour issues, pay little attention to content, attend to the teacher's rather than the pupils' actions and tend to make general comments, evaluate rather than interpret, and use naïve assumptions rather than theory (e.g., McDonald, 2016; Schäfer & Seidel, 2015; Simpson et al., 2018; Sonmez & Hakverdi-Can, 2012). While acknowledging that PSTs cannot be expected to possess advanced professional vision, researchers implicitly use a normative frame of reference of what PSTs should notice to demonstrate skills, which then serves as a long-term goal to achieve. There seems to be an agreement that teachers should be able to attend to pupils' and not only to the teacher's actions, to pay attention to the content in relation to the pupils and use theory to explain the observed. A few studies make target noticing (Stockero & Rupnow, 2017) explicit and indicate what teachers are supposed to notice to demonstrate good noticing. They create expert norms which determine situations in the lesson pertinent to its success. For example, in Wiens et al.'s (2021) study, a specific framework was developed by researchers to measure how the PSTs noticed the nature of pupil-teacher interactions.

Professionals creating the expert norm vary. They may be experienced teachers (McDonald, 2016; Star et al., 2011) or researchers who follow and conduct research, write and review papers, teach at the university, etc., and thus, are presumed to have "acquired integrated knowledge structures" (Schäfer & Seidel, 2015, p. 54). In the field of professional vision, experts may be researchers and teacher educators who specialise in instructional design and educational research (Blomberg et al., 2011; Shäfer & Seidel, 2015) or in subject education (Mitchell & Marin, 2015; Steffensky et al., 2015; Stockero & Rupnow, 2017; Vondrová & Žalská, 2015).

The above opens up an issue of differences between the researcher professional vision and the practitioner professional vision (Lefstein & Snell, 2011), which may adversely impact setting an expert benchmark. To bridge this divide, Roose et al. (2018) suggest involving a range of experts such as academics, teachers, teacher educators, etc. In our study, the expert norm is a product of plural expert professional visions (Lefstein & Snell, 2011) of teacher educators of different fields.

## 2.2 Instructional quality

In line with Litke, et al. (2021), we understand the quality of teaching as "the extent to which classroom instruction consists of structures and practices believed by researchers and practitioners in the field to provide rich learning experiences for students" (p. 1). In their meta-analysis, Seidel and Shavelson (2007) identified effective teaching variables in goal setting, orientation (mobilising pupils' prior knowledge and investigating possible routes towards the goal), execution of learning activities, evaluation of learning processes, and teacher guidance and support. Killen's (2006) quality teaching model includes intellectual quality, relevance (connectedness), a socially supportive learning environment, and recognition of difference. Even within the subject, there is no widely shared agreement on what quality teaching means. For example, Litke et al. (2021) synthesised three frameworks for quality teaching in mathematics into one model of general elements of teaching quality. They pointed to the merits and pitfalls of looking for one model to capture all dimensions of teaching quality.

Another aspect to consider is the type of lesson under consideration. The characteristics of a successful revision lesson will differ from the ones of the lesson in which a new subject matter is introduced. In our study, we restrict ourselves to the latter. Thus, two dimensions of Litke et al.'s (2021) synthesised model are particularly relevant: *Selecting and addressing the content and subject-specific methods* (motivating the content, addressing the content in structured, accurate, and disciplinary correct ways) and *cognitive activation* (practices creating opportunities for pupils' learning, teacher facilitation of pupils' cognitive activities and of metacognitive learning from cognitively activating tasks). This model is supported by subject-specific studies on instructional quality. For example, Steffensky et al. (2015) highlighted the dimension of learning support with cognitive activation and structuring, and Kaiser et al. (2015, p. 374) emphasised "demanding orchestration of teaching the mathematical subject matter, potential for cognitive activation of the learners, individual learning support and classroom management" as prerequisites of quality teaching. Cognitive activation includes metacognitive activities and pupils' self-regulation and independence, applicable across subjects (Perry et al., 2018).

In our research, we mainly build on the conceptions of Schlesinger et al. (2018) and Janík et al. (2019).

Schlesinger et al.'s (2018) framework comes from mathematics education. It includes two subject-specific dimensions of instructional quality. The dimension of subject-related quality comprises dealing with mathematical errors of students, a teacher's mathematical correctness, a teacher's mathematical explanations, mathematical depth of the lesson and support of mathematical competencies. The dimension of teaching-related quality comprises using multiple representations, deliberate practice, appropriate mathematical examples and relevance of mathematics pupils.

Janík et al. (2019), whose model is based on the analysis of lessons across subjects, connect the quality of instruction to its integrity, namely, to "the quality of functional relationships between (1) teaching and learning content, (2) teaching and learning objectives and (3) the activities of a teacher and students" (p. 189). Situations with high integrity are examples of *didactic excellence*, while situations with lower integrity (or even disconnection between the three levels) are examples of *didactic formalism*. They often result from over-focusing on the form (or organization) of teaching at the expense of the content. Janík et al. (2019) present two types of didactic formalism: stolen cognition and concealed cognition.

If the teacher over-reduces the space available for pupils' cognitive work with the content, we are witnesses of *stolen cognition*. Pupils are passive as the content is "remote from their cognitive and motivational states, and the learning environment cannot give them sufficient insight into the content" (Janík et al., 2019, p. 192). Stolen cognition is the result of problems in the selection of content (e.g., it is too distant and demanding), in didactic structuring of content, which leads to a lack of clarity of content representation or in assessment, and work with mistakes which is less formative and does not "support autonomy in learning and cognitive activities well" (ibid. 2019, p. 192).

Situations of *concealed cognition* are "instances of purposeless cognitive activation of students due to their being disconnected from the content" (Janík et al., 2019, p. 185). Pupils are seemingly active and work on the task, but the teaching-learning situation does not enable them to develop a deep understanding of the content. Pupils "miss important elements of content and crucial relationships between them; they 'lose themselves' in content — distort it, make it too easy or notice only its unimportant aspects" (ibid. 2019, p. 197). Janík et al. (2019) posit that concealed cognition might be less obvious than stolen cognition because pupils are usually keen to work on the task, and thus, a teacher does not get warning feedback, and that is the reason why examining the manifestation of didactic formalism in teaching and learning is important.

## 2.3 Research aim

To sum up, only a few studies focus on the compatibility of the views of experts and PSTs or practising teachers when analysing a (video)-lesson (e.g., Blomberg et al., 2011; Mitchell & Marin, 2015; Schäfer & Seidel, 2015; Stockero & Rupnow, 2017; Stürmer et al., 2013). Yet, the alignment between PSTs' views and what experts consider appropriate is important (Schäfer & Seidel, 2015) as it informs teacher educators of where PSTs are on their way to a long-term target (expert-like noticing and knowledge-based reasoning).

As our research concerns the analysis of teaching in different fields, a shared view of quality teaching across these fields was needed. We grounded it in the characteristics of the teaching quality, which we found applicable across our respective fields. We selected the dimensions of content representation, content selection, didactic structuring of content, and assessing and dealing with mistakes (Janík et al., 2019), and a teacher's correctness, content depth of the lesson, multiple representations, and content relevance (Schlessinger et al., 2018).

This paper consists of two related studies. The first aims to develop an expert target, a framework for identifying content-related situations in the lesson deemed important by experts for its instructional quality. The second study aims to determine whether PSTs notice phenomena deemed important by experts and whether their comments are compatible with the experts' ones. The second study compares PSTs' noticing and knowledge-based reasoning against the target developed in the first study.

# 3 Study 1: Identification of expert phenomena

In Study 1, we developed a framework for determining important situations in the lesson in which a new subject matter is introduced. Its bases were the characteristics and indicators for the teaching quality of Janík et al. (2019) and Schlessinger et al. (2018), as given above. This framework was applied to videos of lessons, and phenomena pertinent to their instructional quality (*expert phenomena*) were determined.

## 3.1 Participants

The noticing target was made by six teacher educators (authors) from different educational fields (art, biology, mathematics, English as a second language, and general education) with similar research and work experience. Each educated PSTs at the same university, taught subject and subject education courses and supervised PSTs during school placements. All conducted research in their respective fields and had experience with cross-subject research on noticing. They had 7 to 36 years of experience at the university, and five had experience teaching at primary and/or secondary schools.

## 3.2 Video selection

First, we decided to use videos of whole lessons rather than clips as they provide a broader context of the teaching-learning situations. Second, we agreed to use lessons comprising the introduction of a new subject matter. Lessons focusing on revision tend to be repetitive. In contrast, lessons with a new subject matter introduction include more diverse teaching-learning situations that merit attention and cannot be envisaged fully. The course of such lessons is dependent on pupils' reactions.

Each researcher selected videos of three lessons with which they had experience from their courses, and the whole team evaluated their suitability. The lessons were to be self-contained (with no need for additional context) authentic lessons from Czech schools, with new content (familiar to PSTs) being introduced. It is considered a sign of teacher noticing if a breach of a norm (Dreher et al., 2021) regarding an aspect of instructional quality is noted. Teaching-learning situations bearing signs of didactic formalism (Janík et al., 2019) can be seen as a breach of the norm of quality teaching. Thus, the selected lessons also included examples of stolen and concealed cognition.[1] Our assumption, supported by our experience with using videos of lessons with PSTs in subject education courses, was that such situations might be easier to recognise for PSTs with little teaching experience and will motivate them to comment on such situations and suggest alterations. Finally, it was considered whether the PSTs had had an opportunity to gain sufficient pedagogical content knowledge (Shulman, 1986) during their subject education courses (taught by the authors) to spot and interpret such situations.

One lesson on elementary art education (EAE), elementary social studies (ESS), biology (BI), the English language as a second language (EL), and mathematics (MA) was assigned for pilot reflection to PSTs studying the subject in question. They were given the task: "You will see a video which captures a lesson on [subject]. You can watch the video as many times as you want. Write a reflection; the length is not specified. Write down what you find interesting; what is, in your opinion, important." Their written responses were analysed to see whether they felt motivated to comment on situations connected to the introduction of new knowledge and situations of didactic formalism and whether the lessons made sense to them.

## 3.3 Designing an analytical framework

For each of the five lessons, we proceeded as follows. We watched the lesson individually, and using Janík et al.'s (2019) and Schlessinger et al.'s (2018) frameworks, we distinguished teaching-learning situations in which work with content could be seen and described them. Next, we compared the lists of such situations and created a master list of the ones mentioned by at least half of the team. We met, watched the lesson together, stopped it at moments from the master list, and discussed their importance for the lesson's success and possible interpretations. The discussion was led by the team member within whose expertise the lesson fell (e.g., the mathematics educator led the discussion for the mathematics lessons). It was audiotaped and summarised by the discussion leader. The team provided their written feedback on this summary, considering whether the suggested situations were related to the quality of teaching, whether they agreed on their interpretation, and if PSTs, in their opinion, had enough knowledge to reach such an interpretation. The most frequent argument for excluding the situation was that we agreed it might only be visible to the experts in the respective field and not to a PST with hardly any teaching experience.

[1]Note that it does not mean that the lessons were examples of "bad" teaching; they included situations which can be classified as examples of didactic excellence (Janík et al., 2019) as will be seen in examples below.

Six to eight teaching-learning situations originated for each lesson, and we negotiated them until an agreement on six was reached (to compare responses for different lessons). The description of the lesson's specific situations and the agreed-on interpretation (causes, implications, explanations supported with theoretical notions, alternative actions) is called *an expert phenomenon* here. Next, a categorisation of the expert phenomena was sought, and after multiple discussions, six categories emerged.

In the second validation stage, we checked whether the expert phenomena could be used to analyse PSTs' responses regarding the compatibility of views. Two PSTs' responses from the pilot stage were selected for each lesson and scrutinised for any mention of the expert phenomena. Then, we considered the compatibility of the views depicted in the expert framework and those of the PSTs. Similarly to Stockero and Rupnow (2017), this was not always straightforward. We sometimes agreed with the PST's comment only partially, as the PST stated something plausible about the situation but failed to comment on an aspect which we deemed important. Moreover, we realised that some phenomena were more complex than others and/or spanned more time and included several aspects that the PSTs could mention. We revisited the lists of expert phenomena and distinguished two to four characteristics for each phenomenon. A detailed manual of the expert target originated.

### 3.4 Results: Expert phenomena

The identified expert phenomena fall within six categories. 'Pupils' cognitive activation' captures whether pupils are engaged in gaining knowledge. 'Depth (work with concepts)' depicts whether the concepts are dealt with sufficiently deep and wide. 'Terminology and precision' relates to the correctness and appropriateness of terms related to the content, of the language used, or of definitions of concepts, taking into account the age of pupils in the lesson. 'Mobilising prior knowledge' points to situations where pupils' prior knowledge necessary for the new subject matter is evoked. 'Relevance' concerns situations in which the subject matter is introduced as relevant to pupils' learning or life, pupils are motivated to learn the content, and connections are made to other subjects. 'Representations' describes the situations in which representations of the new content are presented (or not). Situations falling within these categories can be envisaged as examples of didactic excellence or didactic formalism if the norm of quality teaching is breached.

As there are five lessons, each described by six expert phenomena and each such phenomenon is accompanied by 2 to 4 characteristics, it is above the scope of this paper to present all. One expert phenomenon for each lesson is illustrated. They are denoted by the subject acronym, numbers 1 to 6 and a letter implying a specific characteristic. As the identified dimensions of quality teaching overlap and "cannot be completely separated" (Schlesinger et al., 2018, p. 478), our expert phenomena are not uniquely categorised either.

In the situations described in the expert phenomenon EL5 (categorised 'Depth [work with concepts]' and 'Mobilising prior knowledge'), an inductive approach was used for pupils to infer rules for relative pronouns from the text, using their prior knowledge (EL5a). Implementing this approach in grammar teaching has multiple benefits (EL5b) and is considered a sign of quality teaching (e.g., Jean & Simard, 2013). Thus, this lesson includes situations demonstrating the principle of 'selecting appropriate content and subject-specific methods' (Litke et al., 2021). However, using the inductive method is uncommon in Czech foreign language teaching, and PSTs might see it as a breach of the norm. This might lead them to suggest that only the deductive method is appropriate when introducing rules for relative pronouns.

In the situations captured in the expert phenomenon MA6 (categorised as 'Pupils' cognitive activation' and 'Depth [work with concepts]'), the teacher assigned the pupils two potentially cognitively challenging tasks which could be used for the assessment of pupils' understanding of the newly introduced Thales's theorem (MA6a). However, their implementation was not adequate. One task could be answered without much thought, and the other was too difficult for the pupils at this stage of learning to solve within the short time provided by the teacher (MA6b). This situation bears signs of a breach of the principles of 'selecting appropriate content' and 'teacher facilitation of pupils' cognitive activity' (Litke et al., 2021).

During the situations described in the expert phenomenon BI4 (categorised as 'Pupils' cognitive activation' and 'Representations'), multiple representations of products of nature (e.g., shells) were used with a potential for pupils' cognitive activation (BI4a). Teaching with the support of products of nature has been shown beneficial to pupils' understanding of content (e.g., Sugni et al., 2011). However, the pupils in the video could not touch objects, and only some could see them properly (BI4b). While the content and representations were well selected, the subject-specific method was not, and the teacher did not create good opportunities for pupils' learning.

In the situations described in the expert phenomenon EAE5 (categorised as 'Relevance' and 'Mobilizing prior knowledge'), the pupils were asked to reflect in groups on themselves in relation to the others to prepare for the subsequent artistic creation of a portrait. Their ability to think conceptually about human identity was targeted (EAE5a). The teacher did not emphasize the difference between internal and external identity (EAE5b), which remained blurred (Mason & Buschkuehle, 2013). The examples provided by the teacher were too simplistic and instructive (EAE5c). The situation might be seen as an example of stolen cognition. The chosen topic demands pupils to express themselves on sensitive and discrete issues (EAE5d) while sensitively maintaining the possibility of silence.

In the situations described in the expert phenomenon ESS4 (categorised as 'Pupils' cognitive activation' and 'Relevance'), the teacher suggested that pupils organise a campaign against food waste which could potentially lead to their commitment to the content (ESS4a). However, the teacher did not include the pupils in the campaign planning or motivate them (ESS4b). The pupils could have been led to realise that they could assist in food waste prevention (ESS4c), which could serve as a proxy for authentic learning situations that benefit pupils' learning (e.g., Cheng et al., 2019).

The identified expert phenomena make an expert target against which PSTs' noticing and reasoning in Study 2 are compared.

# 4 Study 2: Pre-service teachers' noticing and knowledge-based reasoning

The research questions of Study 2 were the following:
**RQ1:** What expert phenomena identified in Study 1 do PSTs at the end of their university study notice in the lessons?
**RQ2:** How does PSTs' knowledge-based reasoning compare to that of experts?

## 4.1 Participants and the task

The study is situated at a faculty educating teachers in the Czech Republic. Future secondary teachers complete a 3-year bachelor's degree and a 2-year master's degree. The study for elementary teachers consists of a 5-year undivided master's degree. PSTs take subject and subject education courses, pedagogical and psychological courses and undergo school practice placements. After the defence of a master's thesis and the final state examination, they become qualified teachers. There is no induction period.

The PSTs did not attend any courses aimed at the development of noticing. However, they dealt with the concept of quality teaching in their general education and subject education courses, in which they were also encouraged to justify their ideas about teaching and learning. Videos of lessons were sometimes used in their courses to accompany the content.

The PSTs were in their $4^{\text{th}}$ or $5^{\text{th}}$ year (Tab. 1) (age $M = 24$, $SD = 5.2$, 89% female). 20% PSTs had teaching experience ($M = 3$ years, $SD = 2.1$ years). All PSTs in the year group studying the programme participated. An exception is EAE, where 34 PSTs out of 36 invited participated.

**Tab. 1:** Participants

| PSTs | Group | Acronym | $N$ | Total |
|---|---|---|---|---|
| Elementary | Elementary English Language | EEL | 23 | 80 |
| | Elementary Social Studies | ESS | 23 | |
| | Elementary Art Education | EAE | 34 | |
| Secondary | English Language | EL | 25 | 94 |
| | Biology | BI | 43 | |
| | Mathematics | MA | 26 | |

The PSTs were assigned the task in their subject education course: "You will see a video which captures a lesson on [subject]. You can watch the video as many times as you want. Write a reflection; the length is not specified. Write down what you find interesting; what is, in your opinion, important. Do not feel afraid to write your views; there are no correct answers. You will not be assessed according to your reflection." The task was rather open, not to focus PSTs' attention on anything. It was assigned as homework to give them enough time to watch the lesson repeatedly and provide multiple opportunities to notice important situations. Each student watched one lesson.

## 4.2 Data analysis

The data consists of the PSTs' written responses to one lesson. They were scrutinised for any mention of expert phenomena particular to the lesson, and all such mentions made one unit of analysis; each unit concerns one expert phenomenon. One researcher split responses into units (Tab. 2), and the second checked its validity. Any inconsistencies were negotiated. The units ranged from one sentence to several paragraphs.

**Tab. 2:** Number of units

| Group | EEL | ESS | EAE | EL | BI | MA | **TOTAL** |
|---|---|---|---|---|---|---|---|
| N | 50 | 75 | 89 | 59 | 141 | 90 | **504** |

The units were coded using the coding framework in Tab. 3. The levels of description and evaluation do not attest to any knowledge-based reasoning. The level of theorising is considered more expert-like than the level of explanation. As the lessons included breaches of a norm which naturally leads to suggesting alternatives, we also coded the units for their presence (Alteration).

**Tab. 3:** Coding framework

| | |
|---|---|
| **Description** | Recounting |
| **Evaluation** | Subjective judgment without explaining |
| **Explanation** | Layman (naïve) explanation or explanation based on one's experience as a pupil or a teacher |
| **Theorising** | Generalisation with theory |

We coded the same four reflections independently, then met, negotiated our agreement, and modified the coding manual accordingly. Next, the reflections of each group were coded by two researchers, one of them being a specialist in the particular field (but her role was not prioritised over the role of the non-specialist). The researchers worked independently (inter-rater reliability found as per cent agreement was from 80.2% to 91.0%) and negotiated any differences until an agreement was reached.

Next, each unit coded Evaluation, Explanation or Theorising was assigned a value of 0/1 for each expert phenomenon's characteristics, depending on whether the PST's comment was consistent with the experts' one. The same pairs of researchers as in the first stage coded the units (inter-rater reliability ranged from 85.1% to 89.1%) and negotiated any differences until an agreement was reached. The decision about each unit was *Match* (the unit received all 1s), *Limited match* (the unit received more or the same number of 1s than 0s), or *No match* (in other cases). Thus, the unit was assigned *No match* if the PST did not comment on more than half of the characteristics of the expert phenomenon and/or their interpretation of the phenomenon was inconsistent with the experts' view.

Finally, units coded Alteration were assigned *Match/No match* with the experts' view. At this stage, the role of the specialist in the field was important, and the final decision was hers.

Examples of PSTs' comments and their coding are in Section 4.3.

## 4.3 Results

On average, the PSTs commented on half of the expert phenomena (Tab. 4, Fig. 1).

**Tab. 4:** Number of mentioned expert phenomena per PST

| | EEL | ESS | EAE | EL | BI | MA | **TOTAL** |
|---|---|---|---|---|---|---|---|
| Mean | 2.2 | 3.3 | 2.6 | 2.4 | 3.3 | 3.5 | **2.9** |
| Median | 2.0 | 3.0 | 3.0 | 2.0 | 3.0 | 3.5 | **3.0** |

Nearly half of the units show no knowledge-based reasoning (Description 18.5%, Evaluation 27.4%). In the rest, explanation without theory prevails (Explanation 34.5%, Theorising 19.6%); see Tab. 5.

**Tab. 5:** Units according to knowledge-based reasoning per group (%)

| | Description | Evaluation | Explanation | Theorising |
|---|---|---|---|---|
| EEL | 26.0 | 10.0 | 28.0 | 36.0 |
| ESS | 22.7 | 24.0 | 28.0 | 25.3 |
| EAE | 15.7 | 30.3 | 44.9 | 9.0 |
| EL | 13.6 | 6.8 | 33.9 | 45.8 |
| BI | 22.7 | 34.8 | 29.8 | 12.8 |
| MA | 10.0 | 38.9 | 41.1 | 10.0 |
| **Total** | **18.5** | **27.4** | **34.5** | **19.6** |

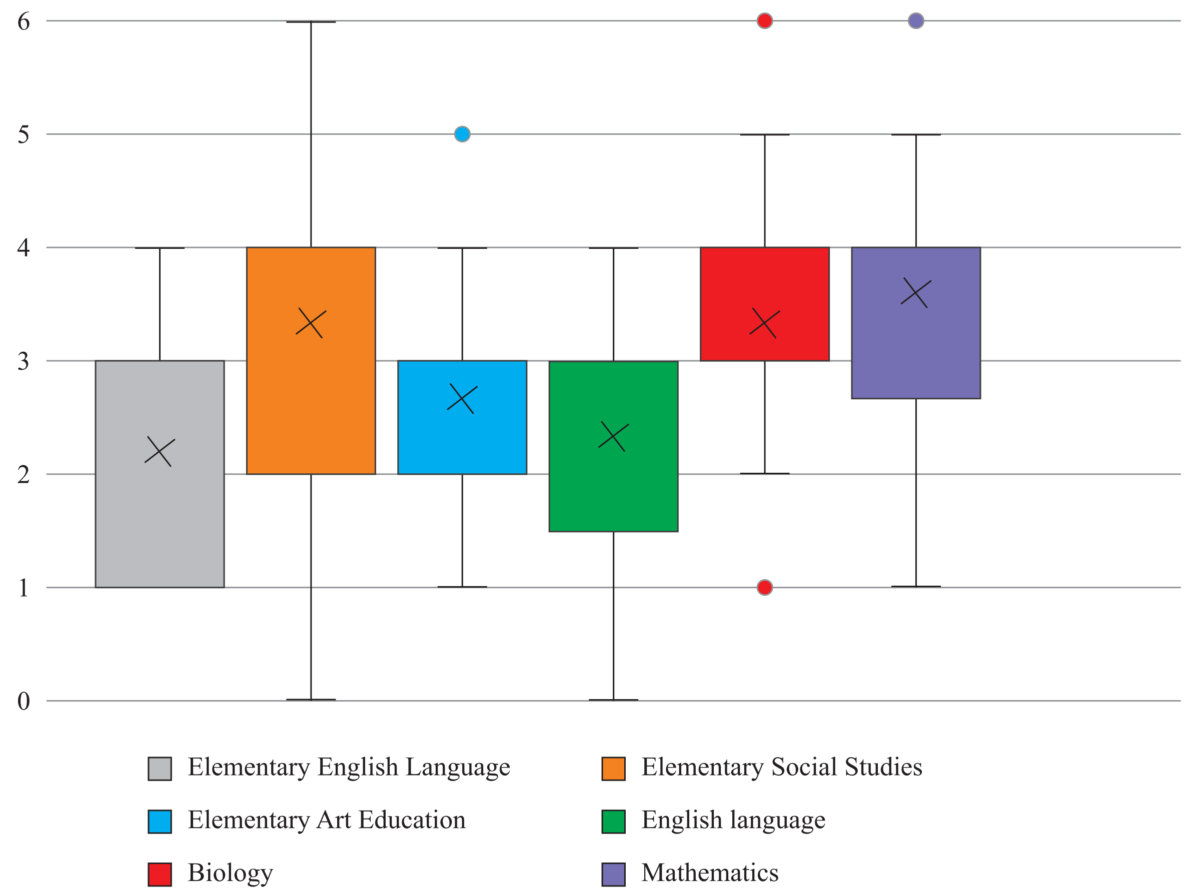



**Fig. 1:** Number of mentioned expert phenomena per PST (× depicts the mean, the horizontal segment is the median, and the plots depict the maximum, minimum, and upper and lower quartiles)

Our focus was on the content of the comments coded Evaluation and above regarding the compatibility of views of the PSTs and the experts. For example, a comment categorised as MA6 (see Section 3.4) "The teacher asked questions related to Thales's theorem which made the pupils think." was assigned 0 for both MA6a and MA6b. The decision was *No match* as the PST did not mention any didactic potential of the tasks, and the experts doubted that the pupils were cognitively activated. The comment "The pupils got two yes-no questions from the teacher which they should answer. In this way, she found out that the pupils understood the subject matter." received 1 for MA6a and 0 for MA6b as the missed learning opportunity was not recognised (*Limited match*).

The comment coded EL5 (see Section 3.4) "The grammar (relative pronouns) was taught inductively. From specific instances, they [the pupils] arrived at the general principle. It is important that the pupils came to a conclusion themselves." was assigned 1 for both EL5a and EL5b and the final decision was *Match*. The comment "The grammar was taught inductively, at first, the pupils worked with the relative pronouns [...] and then formulated the rule for single relative pronouns." was assigned 1 for recognising an inductive approach (EL5a) and 0 as no benefit was mentioned (EL5b). The final decision was *Limited match*.

For the whole sample, comments in 29.6% of the units agreed with the experts' views, 35.4% agreed partially, and 35.0% were in disagreement (Tab. 6).

**Tab. 6:** Compatibility of the PSTs' and experts' views (%)

| | *Match* | *Limited match* | *No match* |
|---|---|---|---|
| EEL | 70.3 | 27.0 | 2.7 |
| ESS | 22.4 | 34.5 | 43.1 |
| EAE | 21.3 | 50.7 | 28.0 |
| EL | 78.4 | 21.6 | 0.0 |
| BI | 9.4 | 37.4 | 53.3 |
| MA | 19.8 | 32.1 | 48.2 |
| **TOTAL** | **29.6** | **35.4** | **35.0** |

Note: 100% is the number of units categorised as Evaluation and above.

Next, we related the compatibility of PSTs' and experts' views and the levels of knowledge-based reasoning. Tab. 7 shows that most *No match* units were on the evaluation level, and the biggest share of *Match* was on the theorising level. Comments coded *No match/Evaluation* ($N = 70$) are the most inadequate as PSTs do not provide any explanation, and their subjective judgment does not concern the selected characteristics of the expert phenomenon or is contrary to the experts' view. On the other hand, the most advanced comments (*Match/Theorising*, $N = 45$) are signs of developed professional vision. Most of the PSTs' comments were between the two poles.

**Tab. 7:** Compatibility according to the level of knowledge-based reasoning (%)

| | *Match* | *Limited match* | *No match* |
|---|---|---|---|
| Evaluation | 14.5 | 34.8 | 50.7 |
| Explanation | 32.6 | 39.0 | 28.5 |
| Theorising | 45.5 | 30.3 | 24.2 |
| **TOTAL** | **29.6** | **35.4** | **35.0** |

The PSTs' responses included 112 alterations to the teaching in the lessons (Fig. 2).

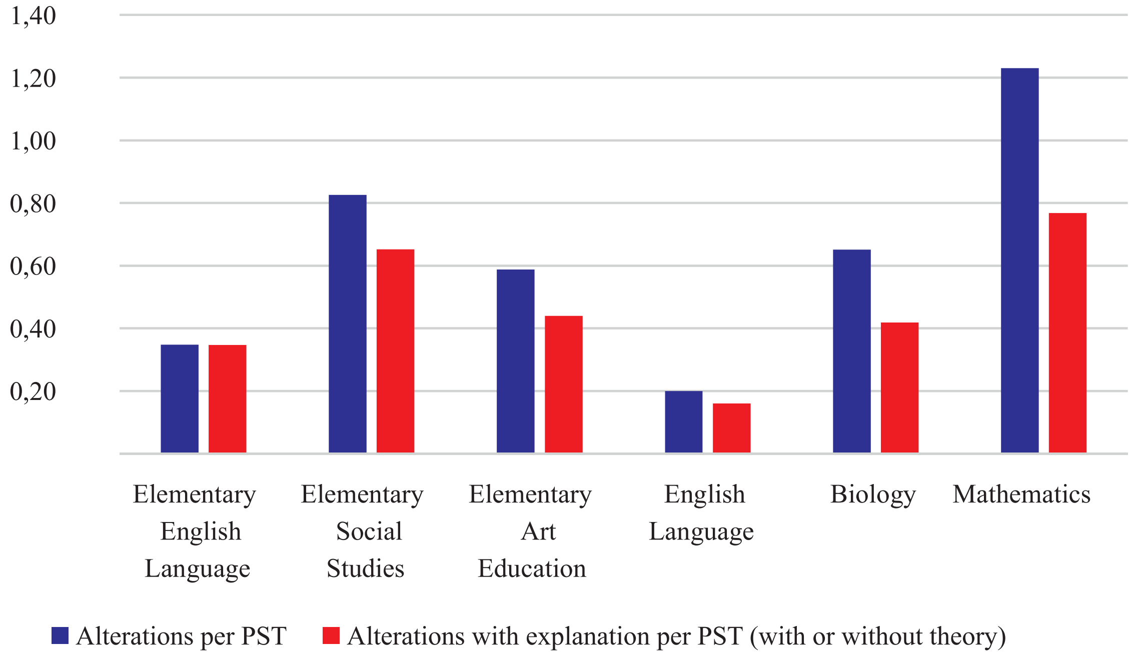



**Fig. 2:** Average number of alterations per PST

The lowest level alterations were accompanied by subjective judgment only (28.6%). Most suggested alterations (49.1%) were explanative. For example, a PST suggested for the ESS lesson: "I find the example of the amount of food wasted by people in Asia and Europe and North America valuable. However, judging by some pupils' statements, I am not sure they understood that the amount of food wasted by European people also concerns themselves, so if I were the teacher, I would make them aware of this fact."

About 22.3% of the alterations were supported by theory. For example, for the EL lesson, a PST wrote: "The teacher sometimes switched to the mother tongue. Very often, it wasn't necessary to use Czech because it wasn't anything the children wouldn't understand. When she said 'To jsem zvědavá! Výborně! Buďte opatrní.',[2] she could have said it in English (or both – English and Czech) to let the pupils acquire the second language. This subconscious process would result in permanent knowledge according to Krashen's Theory of Second Language Acquisition."

Table 8 shows that the PSTs mostly suggested alterations coded as *Match* (82.1%), but it does not necessarily mean that the experts suggested the same. For example, the pupils were asked to formulate Thales's theorem towards the end of the MA lesson. A PST appreciated this task but suggested that the theorem could have been written in "a more mathematical" way by the teacher. This comment was coded *Match*, as the experts would welcome it if the theorem were correctly written mathematically. However,

[2] 'I want to see this! Very well! Be careful!'

**Tab. 8:** Alignment of PSTs' alterations with experts (%)

| | *Match* | *No match* |
|---|---|---|
| EEL | 87.5 | 12.5 |
| ESS | 79.0 | 21.1 |
| EAE | 80.0 | 20.0 |
| EL | 60.0 | 40.0 |
| BI | 89.3 | 10.7 |
| MA | 81.3 | 18.8 |
| **TOTAL** | **82.1** | **17.9** |

they saw a more pressing alteration in this part of the lesson; the task was above the pupils' ability at this stage of learning, and more time should be devoted to letting pupils formulate the theorem in their own words.

According to the experts, *No match* (17.9%) alterations would not enhance teaching. An example is a comment on the BI lesson: "I do not think this way of teaching should be used to introduce a new subject matter. A table is an excellent tool which enables sorting out information. Still, I would use such type of lesson only for the revision of the subject matter." According to the experts, sorting out data through a table can be successfully used for the inductive introduction of subject matter. Another example is a PST's suggestion "to concentrate on only one activity leading to Thales's theorem, not two". From the PST's analysis, it was clear that he does not understand that the two activities are not the same but differ in the direction. It does not agree with the experts' view as the two "activities" in the lesson represent two directions of implication which make Thales's theorem.[3]

## 5 Discussion and implications

It is documented in the literature what PSTs pay attention to and what they neglect in a video lesson. It is less clear how their noticing and reasoning relate to that of experts. Based on the frameworks of instructional quality (Janík et al., 2019; Kaiser et al., 2015; Litke et al., 2021; Schlesinger et al., 2018; Seidel & Shavelson, 2007), we determined six dimensions which we found applicable in selected lessons across subjects. In Study 1, the experts distinguished and interpreted the situations deemed important for pupils' learning of new content to develop a target against which the PSTs' responses were compared in Study 2. Our literature search confirmed that studies using such targets are rare (McDonald, 2016; Mitchel & Marin, 2015; Steffensky et al., 2015; Stockero & Rupnow, 2017; Vondrová & Žalská, 2015).

We showed that PSTs at the end of their university study found it difficult to notice content-related situations deemed important by the experts, as they mentioned half of the expert phenomena. The same conclusion was reached for PSTs' ability to notice generic aspects pertinent to instructional quality (Schäfer & Seidel, 2015). Subject-focused research on expert-like noticing seems to exist for mathematics only. In Stockero and Rupnow's (2017) study, PSTs noticed one-third of the events determined by experts as examples of Mathematically Significant Pedagogical Opportunity to Build on Student Thinking. PSTs' mean score in Star et al.'s (2011) study for what they call important moments was 53% (ours was 48%). Little attention to the subject-related expert phenomena was found in a study with mathematics PSTs at the end of university with the same educational experience as our MA group (Vondrová & Žalská, 2015); the median was 2, while the expert value was 7. The results in the present study are better, which can be attributed to the way the expert target originated. Mathematics educators made it in Vondrová and Žalská's study, while experts from five fields were involved in the present study and the phenomena deemed important by the experts in mathematics education but not visible to the experts outside the field were excluded.

About half of the comments of our sample bore no sign of knowledge-based reasoning. The PSTs described the events and/or made personal judgments about them. In the other half, an explanation based on their experience as pupils or naïve assumptions prevailed. Concepts from learning and teaching theories introduced during their university studies were only present in 20% of comments. The PSTs in Schäfer and Seidel's (2015) study also struggled when reasoning about important events. For goal clarity, nearly 73% of their comments were naïve assumptions with judgmental character, and only 27% included the use of professional knowledge. For learning climate, it was 89% against 11%.

Another focus of our study was on the agreement between the PSTs and experts. The PSTs' comments fully aligned with those of the experts in less than a third of the cases and partially in another third.

[3]In Czech textbooks, Thales' theorem is formulated as an equivalence.

The PSTs who provided the rest of the comments could not apply their knowledge outside the university course. This supports the literature showing that PSTs struggle to apply theoretical notions in practice (e.g., Hammerness et al., 2002).

Two of the few studies comparing the PSTs' and experts' views available to us reached similar results. In Schäfer and Seidel's (2015) study, the PSTs' reasoning matched the experts' in about a third of their statements for goal clarity and a quarter for learning climate. Blomberg et al. (2011) found for PSTs of different subjects that the average agreement with the experts' views was 31%.

Determining what is important in a lesson is not trivial for PSTs. Moreover, we must remember that PSTs' "acquired professional knowledge is not yet very elaborated and still is determined by naïve judgments and subjective theories" (Schäfer & Seidel, 2015, p. 36). Our study supports the body of research calling for developing PSTs' professional vision in both its aspects (noticing and knowledge-based reasoning), e.g., in video interventions supported by scaffolding frameworks (e.g., Santagata, 2011; Stockero & Rupnow, 2017). As the PSTs struggled with noticing situations deemed important by the experts, we suggest that university courses target the content dimensions of instructional quality and how they manifest themselves in lessons. The PSTs would also benefit from tasks on reasoning about the consequences of events and suggesting alterations. They should be given multiple opportunities to apply theoretical knowledge to concrete lessons.

## 6 Limitations and conclusions

Our conclusions should be considered in light of a limited sample. It did not allow us to draw any conclusions regarding similarities and differences between subject groups which the tables presenting results seem to suggest. In some cases, the number of units is low, and no conclusions are possible about differences (see, e.g., the number of alterations suggested by all the PSTs or the number of comments coded Evaluating, Theorising, etc.). Some differences, though, seem to be present. For example, concerning the two English language groups (EEL and EL), they used theorising most often (Tab. 5), and the vast majority of their comments were in agreement with the experts' views (Tab. 6). This observation may be a result of the selected lesson. The EEL lesson might provide more opportunities (and more recognisable) than the other lessons used in our study for PSTs to apply a theory they learned. In such a case, they could be assigned another lesson for reflection (with a different topic, different teaching-learning situations, etc.). Then the analysis of the two lessons could be compared. However, such a task would be enormous (if not impossible), given the number of characteristics present in a single lesson which could be varied. Another reason for the diverging results of the EEL/EL groups might be that they learned more theory (or in a more appropriate way, enabling them to grasp it better) than the other PSTs in our study in their subject education courses. Yet another reason might lie in the sample itself – this group of PSTs might have more developed knowledge-based reasoning than the other groups within our sample. Indeed, Simpson and Vondrová (2019) found differences in noticing skills between two samples of mathematics PSTs with the same background and in the same stage of their university education.

Thus, at this stage of analysis, no conclusions can be made about the differences between subject groups. However, in a future study, the analysis could delve deeper into the content of the PSTs' comments to see the characteristics of those expert phenomena in which the PSTs were (or were not) able to apply theory and agreed (or did not agree) with the experts' views. This goes beyond this paper and would merit a separate paper for each subject group.

Another limitation of our study is the type of task we used – it was formulated openly. A future study could use a more specific task to focus PSTs' attention and, moreover, ask for justifications and alterations, which we did not do. Assigning the task as homework might have enabled the PSTs not to make their best effort. However, the diversity and length of their responses suggest that it was not the case.

Another limitation concerns the noticing target. It originated through a comparative judgment whose validity is based on the collective expertise of a community of experts from different fields. Still, it remains subjective and might be skewed towards the researcher professional vision. In a future study, experienced teachers could participate in the expert analysis (Lefstein & Snell, 2011; Roose et al., 2018). No experts outside the research team validated the noticing target, which is another limitation (although not uncommon, e.g., in Mitchel and Marin's (2015) study, both researchers were master raters). Dreher et al. (2021) point out that any norm for teacher noticing is necessarily culture-specific; thus, our conclusions must be read in Western culture.

The question of visibility of situations comes to the fore as our sample did not see the same lesson. Similarly to Blomberg et al. (2011), we tried to account for this by balancing the input from an expert in the field, able to spot hidden features of the lesson, with the inputs from researchers from different fields. By an elaborate way of selecting the phenomena for each lesson, we strove to ensure that they were similarly observable. Moreover, all the selected phenomena allowed for higher levels of reasoning and, thus, presented comparable conditions for PSTs to manifest their knowledge-based reasoning.

To sum up, our study brought new insights into the understudied area. Unlike prevailing research focusing on what PSTs attend to and how they interpret it, we also investigated the content of PSTs' comments and whether it aligns with what experts accept. Most PSTs in our sample about to start teaching could not discern situations important for deep work with content in the lesson and apply their theoretical knowledge for interpreting them. The ability to do so is necessary for their future careers as teachers, and thus, we need to know how their noticing and reasoning relate to those of experts. More studies are needed to validate our results with PSTs of various subjects and different lessons (or clips).

## References


Blomberg, G., Stürmer, K., & Seidel, T. (2011). How pre-service teachers observe teaching on video: Effects of viewers' teaching subjects and the subject of the video. *Teaching and Teacher Education*, *27*(7), 1131–1140. https://doi.org/10.1016/J.Tate.2011.04.008

Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies. *Zeitschrift für Psychologie*, *223*(1), 3–13. https://doi.org/10.1027/2151-2604/a000194

Cheng, S.-C., Hwang, G.-J., & Chen, C.-H. (2019). From reflective observation to active learning: A mobile experiential learning approach for environmental science education. *British Journal of Educational Technology*, *50*(5), 2251–2270. https://doi-org.ezproxy.is.cuni.cz/10.1111/bjet.12845

Dreher, A., Lindmeier, A., Feltes, P., Wang, T.-Y., & Hsieh, F. J. (2021). Do cultural norms influence how Teacher noticing is studied in different cultural contexts? A focus on expert norms of responding to students' mathematical thinking. *ZDM, 53*, 165–179. https://doi.org/10.1007/s11858-020-01197-z

Goodwin, C. (1994). Professional vision. *American Anthropologist*, *96*(3), 606–633.

Hammerness, K., Darling-Hammond, L., & Shulman, L. (2002). Toward expert thinking: How curriculum case writing prompts the development of theory-based professional knowledge in student teachers. *Teaching Education*, *13*(2), 219–243. https://doi.org/10.1080/1047621022000007594

Janík, T., Slavík, J., Najvar, P., & Janíková, M. (2019). Shedding the content: Semantics of teaching burdened by didactic formalisms. *Journal of Curriculum Studies*, *51*(2), 185–201. https://doi.org/10.1080/00220272.2018.1552719

Jean, G., & Simard, D. (2013). Deductive versus inductive grammar instruction: Investigating possible relationships between gains, preferences and learning styles. *System*, *41*(4), 1023–1042. https://doi.org/10.1016/j.system.2013.10.008

Kaiser, G., Busse, A., Hoth, J., König, J., & Blömeke, S. (2015). About the complexities of video-based assessments: Theoretical and methodological approaches to overcoming shortcomings of research on teachers' competence. *International Journal of Science and Mathematics Education*, *13*(2), 369–387. https://doi.org/10.1007/s10763-015-9616-7

Killen, R. (2006). *Effective teaching strategies. Lessons from research and practice*. Thompson Social Science Press.

Lefstein, A., & Snell, J. (2011). Professional vision and the politics of teacher learning. *Teaching and Teacher Education*, *27*(3), 505–514. https://doi.org/10.1016/j.tate.2010.10.004

Litke, E., Boston, M., & Walkowiak, T. A. (2021). Affordances and constraints of mathematics-specific observation frameworks and general elements of teaching quality. *Studies in Educational Evaluation*, *68*, 100956. https://doi.org/10.1016/j.stueduc.2020.100956

Mason, R., & Buschkuehle, C. P. (Eds.) (2013). *Images and identity: Educating citizenship through digital arts.* Intellect Ltd.

McDonald, S. P. (2016). The transparent and the invisible in professional pedagogical vision for science teaching. *School Science and Mathematics*, *116*(2), 95–103. https://doi.org/10.1111/Ssm.12156

Mitchell, R. N., & Marin, K. A. (2015). Examining the use of a structured analysis framework to support prospective Teacher noticing. *Journal of Mathematics Teacher Education*, *18*(6), 551–575. https://doi.org/10.1007/s10857-014-9294-3

Perry, J., Lundie, D., & Golder, G. (2018). Metacognition in schools: what does the literature suggest about the effectiveness of teaching metacognition in schools? *Educational Review*, *71*(4), 483–500. https://doi.org/10.1080/00131911.2018.1441127

Roose, I., Goossens, M., Vanderlinde, R., Vantieghem, W., & Van Avermaet, P. (2018). Measuring professional vision of inclusive classrooms in secondary education through video-based comparative judgement: An expert study. *Studies in Educational Evaluation*, *56*, 71–84. https://doi.org/10.1016/j.stueduc.2017.11.007

Santagata, R. (2011). From teacher noticing to a framework for analyzing and improving classroom lessons. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 182–198). Taylor and Francis.

Schäfer, S., & Seidel, T. (2015). Noticing and reasoning of teaching and learning components by pre-service teachers. *Journal for Educational Research Online/Journal Für Bildungsforschung Online*, *7*(2), 34–58.

Schlesinger, L., Jentsch, A., Kaiser, G., König, J., & Blömeke, S. (2018). Subject-specific characteristics of instructional quality in mathematics education. *ZDM*, *50*(3), 475–490. https://doi.org/10.1007/s11858-018-0917-5

Seidel, T., & Shavelson, F. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, *77*(4), 454–499.

Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, *60*(1), 20–37. https://doi.org/10.1177/0022487108328155

Sherin, M. G., Russ, R. S., & Colestock, A. A. (2011). Assessing mathematics teachers' in-the-moment noticing. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 79–94). Taylor and Francis.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14. https://doi.org/10.3102/0013189x015002004

Simpson, A., & Vondrová, N. (2019). Developing pre-service teachers' professional vision with video interventions: a divergent replication. *Journal of Education for Teaching*, *45*(5), 567-584. https://doi.org/10.1080/02607476.2019.1674563

Simpson, A., Vondrová, N., & Žalská, J. (2018). Sources of shifts in pre-service teachers' patterns of attention: the roles of teaching experience and of observational experience. *Journal of Mathematics Teacher Education*, *21*(6), 607–630. https://doi.org/10.1007/s10857-017-9370-6

Sonmez, D., & Hakverdi-Can, M. (2012). Videos as an instructional tool in pre-service science teacher education. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, *46*, 141–158.

Star, J. R., Lynch, K., & Perova, N. (2011). Using video to improve preservice mathematics teachers' abilities to attend to classroom features: A replication study In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 117–133). Taylor and Francis.

Steffensky, M., Gold, B., Holdynski, M., & Möller, K. (2015). Professional vision of classroom management and learning support in science classrooms – Does professional vision differ across general and content-specific classroom interactions? *International Journal of Science and Mathematics Education*, *13*(2), 351–368. https://doi.org/10.1007/S10763-014-9607-0

Stockero, S. L. (2008). Using a video-based curriculum to develop a reflective stance in prospective mathematics teachers. *Journal of Mathematics Teacher Education*, *11*(5), 373–394. https://doi.org/10.1007/s10857-008-9079-7

Stockero, S. L., & Rupnow, R. L. (2017). Measuring noticing within complex mathematics classroom interactions. In E. O. Schack, M. H. Fisher, & J. A. Wilhelm (Eds.), *Teacher noticing: Bridging and broadening perspectives, contexts and frameworks* (pp. 281–301). Springer.

Stürmer, K., Könings, K. D., & Seidel, T. (2013). Declarative knowledge and professional vision in teacher education: Effect of courses in teaching and learning. *British Journal of Educational Psychology*, *83*(3), 467–483. https://doi.org/10.1111/J.2044-8279.2012.02075.X

Sugni, M., Barbaglio, A., Bonasoro, F., Gioria, M., Fasano, P., & Pasini, M. E. (2011). The role of models in science: A multicomprehensive experience with the sea urchin Paracentrotus lividus. *Procedia-Social and Behavioral Sciences*, 93, 1404–1408.

Sun, J., & van Es, E. A. (2015). An exploratory study of the influence that analyzing teaching has on preservice teachers' classroom practice. *Journal of Teacher Education*, *66*(3), 201–214. https://doi.org/10.1177/0022487115574103

van Es, E. A., & Sherin, M. G. (2021). Expanding on prior conceptualizations of teacher noticing. *ZDM – Mathematics Education*, *53*, 17–27. https://doi.org/10.1007/s11858-020-01211-4

Vondrová, N., & Žalská, J. (2015). Ability to notice mathematics specific phenomena: What exactly do student teachers attend to? *Orbis scholae*, *9*(2), 77–101. https://doi.org/10.14712/23363177.2015.81

Wiens, P. D., LoCasale-Crouch, J., Cash, A. H., & Romo Escudero, F. (2021). Preservice teachers' skills to identify effective teaching interactions: Does it relate to their ability to implement them? *Journal of Teacher Education*, *72*(2), 180–194. https://doi.org/10.1177/0022487120910692

## Scientia in educatione

*Vědecký recenzovaný časopis pro oborové didaktiky*
*přírodovědných předmětů a matematiky*
*Scientific Journal for Science and Mathematics Educational Research*

Vydává nakladatelství Karolinum – http://www.scied.cz





**Adresa redakce**
Pedagogická fakulta, Univerzita Karlova, Magdalény Rettigové 4, 116 39 Praha 1
e-mail: scied@pedf.cuni.cz

Pokyny pro autory jsou uvedeny na
http://ojs.pedf.cuni.cz/index.php/scied/about/submissions#authorGuidelines.

Sazbu v systému LaTeX zpracoval Ing. Miloš Brejcha, Vydavatelský servis, Plzeň.
Logo navrhl Ivan Špirk.
Redaktorka a jazyková korektorka Mgr. Zdeňka Janušová