

UNIVERZITA KARLOVA V PRAZE

Fakulta tělesné výchovy a sportu

**Příspěvek k diagnostice florbalových dovedností hráčů
mladšího školního věku**

AUTOREFERÁT DISERTAČNÍ PRÁCE

Vedoucí disertační práce:

Doc. PhDr. Josef Dovalil, CSc.

Vypracovala:

Mgr. Zuzana Dragounová

Praha 2018

Abstrakt

Název: Příspěvek k diagnostice florbalových dovedností hráčů mladšího školního věku.

Cíle: Hlavním cílem disertační práce je návrh standardizovaného diagnostického nástroje, posuzovací škály, určené pro diagnostiku florbalových dovedností hráčů florbalu mladšího školního věku, který bude vypovídat o úrovni testovaných florbalových dovedností, která je nezbytná pro budoucí herní výkon.

Metody: Disertační práce byla koncipována jako empirický kvantitativně orientovaný výzkum využívající metodu odborného posuzování a techniku škálování. Pro sestrojení posuzovací škály Guttmanova typu byla použita teorie položkových odpovědí, konkrétně Raschův model. K výzkumu byly využity poznatky z konstrukce perfektní škály pro diagnostiku motorických dovedností sestrojené Čepičkou (2003, 2005), doporučené postupy pro standardizaci motorických testů od autorů Štochl a Musálek (2009), expertní analýza pro posouzení obsahové validity podle Lawshe (1971), ztrátové fit funkce pro určení shody modelu s daty, koeficient KR – 20 pro výpočet reliability, Fleissův kappa koeficient pro zjištění shody posuzovatelů a analýza hlavních komponent pro určení unidimenzionality škály.

Výsledky: Sestrojená a vědecky standardizovaná posuzovací škála určená pro diagnostiku florbalových dovedností hráčů florbalu mladšího školního věku obsahuje devět položek. Postup při konstrukci škály by měl být přínosný pro další možnosti tvorby motorických škál v ostatních sportovních specializacích a kinantropologii.

Klíčová slova: florbal, motorická škála, teorie položkových odpovědí, Raschův model, Guttmanova škála, standardizace, expertní analýza, validita, reliability, Fleissovo kappa, analýza hlavních komponent standardizovaných reziduí, unidimenzionalita.

Abstract

Title: Contribution to the diagnostics of floorball skills of floorball players of younger school age (6 – 12 years).

Objectives: The main goal of the dissertation is to design a standardized diagnostic tool, a rating scale designed to diagnose floorball skills of floorball players of younger school age, which will predicate the level of tested floorball skills that is necessary for future game performance.

Methods: The dissertation employed empirical quantitative research methodology, using the method of expert assessment and scaling technique. For the construction of the Guttman-type assessment scale, the item response theory was applied, namely Rasch model. The methodology employed the technique of the construction of a “perfect scale for motor skills diagnostics” developed by Čepička (2003, 2005); recommended procedures for standardization of motor tests by authors Štochl & Musálek (2009); expert analysis for content validity assessment according to Lawshe (1971); fit functions to determine the fit of the data model; KR – 20 coefficient for the reliability calculation; Fleiss' kappa coefficient for determining the agreement of the raters and the principal component analysis of residuals to determine the unidimensionality of the scale.

Results: The developed and scientifically standardized rating scale designed to diagnose the floorball skills of floorball players of younger school age contains nine items. The process of the designing of the scale should be beneficial for the design of motor tests in other sports specialties and Kinanthropology.

Keywords: floorball, motor scales, item response theory, Rasch model, Guttman scale, standardization, expert analysis, validity, reliability, Fleiss' kappa, principal component analysis of residuals, unidimensionality.

Obsah

1.	ÚVOD	5
2.	TEORETICKÝ ZÁKLAD ŘEŠENÍ	7
3.	CÍL PRÁCE A ÚKOLY PRÁCE	10
4.	METODOLOGIE VÝZKUMU	12
4.1	Výzkumný soubor.....	12
4.2	Výzkumné metody.....	13
4.2.1	<i>Postup při konstrukci motorické škály.....</i>	13
4.2.2	<i>Posouzení validity, reliability, objektivit a unidimensionalit škály</i>	14
5.	VÝSLEDKY	20
5.1	Výsledky expertní analýzy	20
5.2	Výsledky pilotního testování	24
5.3	Sestavení škály	25
5.3.1	<i>1. fáze – vyřazení nevyhovujících položek</i>	25
5.3.2	<i>2. fáze – kvantifikace obtížnosti jednotlivých položek</i>	28
5.3.3	<i>3. fáze – zhodnocení rozložení škálových hodnot</i>	29
5.4	Standardizace škály – validita	31
5.5	Standardizace škály – reliabilita	33
5.6	Standardizace škály – objektivita	34
6.	DISKUSE.....	35
7.	ZÁVĚR	38
8.	LITERATURA	40

1. ÚVOD

Řízený tréninkový proces se musí zabývat efekty, které očekáváme. Informace o změnách, k nimž v důsledku tréninkového procesu dochází, má poskytnout kontrola trénovanosti. Plní tak nezastupitelnou úlohu zpětné vazby. Tyto informace o tréninku absolvovaném v uplynulém časovém období se stávají oporou k úvahám o dalším postupu – zda pokračovat v plánovaném tréninku nebo naopak přistoupit k určitým korekcím (Dovalil a kol., 2012).

Řada sportovních specializací se snaží používat různé diagnostické prostředky za účelem kontroly tréninkového procesu, ve velké většině se však jedná o pouhou zkušeností generované přístupy. Stejná situace je v současné době i ve florbalu. Současné poznatky pedagogiky a didaktiky přitom nabízejí aplikaci objektivnějších přístupů.

Ve florbalu, vzhledem k jeho krátké historii, doposud neexistují pro děti a mládež žádné plně standardizované diagnostické nástroje (testy, posuzovací škály), které by umožnily trenérům objektivně hodnotit úroveň osvojených florbalových dovedností.

Předkládaná studie se zabývá konstrukcí standardizovaného diagnostického nástroje, který bude vypovídat o úrovni testovaných florbalových dovedností, která je nezbytná pro budoucí herní výkon. Ověřený diagnostický nástroj, vytvořený pro hodnocení florbalových dovedností hráčů florbalu, je určen pro kategorii mladšího školního věku.

Mladší školní věk lze rozdělit do dvou období: dětství a prepubescence. První období mladšího školního věku je charakteristické nižší úrovní kvality pohybu, dětská motorika postrádá úspornost a každá činnost je prováděna s množstvím dalších přídatných pohybů. Pozdější prepubescentní období je naopak možné označit za etapu s dobrou charakteristikou kvality pohybů a tedy jako příznivé období pro motorický vývoj a učení se novým dovednostem (Perič, 2012). Pro vytvoření diagnostického nástroje florbalových dovedností jsme se rozhodli pro kategorii mladšího školního věku právě z důvodu nestejně úrovně motorického vývoje, který vyžaduje rozdílné nároky na obsah a obtížnost testů a posuzovacích škál.

Příklady a popisy testů pohybových dovedností a posuzovacích škál nalezneme například v publikacích Měkota a Blahuš (1983), Měkota a Cuberek (2007), nebo

u zahraničních autorů Baumgartner (2003), Knudson a Morrison (2002), Morrow (1998) a Thomas (2005). Posuzovacími škálami v kinantropologii se u nás zabývali Jansa (1989) nebo Čepička (2003). Aplikací výzkumných metod v kinantropologii se v posledních letech věnoval například Zháněl (2014).

Po prostudování diagnostických nástrojů určených k testování pohybových dovedností jsme se rozhodli pro účely naší práce zkonstruovat motorickou škálu. Pro konstrukci škály jsme použili Guttmanův model perfektní škály, který jsme sestavili prostřednictvím Raschovy analýzy. Raschův model je jedním ze základních modelů teorie položkových odpovědí a nabízí se jako vhodný nástroj pro konstrukci perfektní škály (Andrich 1988, Čepička 2003). K výzkumu jsme použili poznatků z konstrukce perfektní škály pro diagnostiku motorických dovedností sestavenou Čepičkou (2003, 2005), doporučené postupy pro standardizaci motorických testů od autorů Štochl a Musálek (2009), expertní analýzu pro posouzení obsahové validity podle Lawshe (1975), ztrátové fit funkce pro určení shody modelu s daty (Blahuš 1996, Linacre 2017, Smith 1998), Fleissovo kappa pro zjištění shody posuzovatelů (Fleiss, 1971), koeficient KR – 20 pro určení reliability škály (Cortina, 1993, Cronbach 1951, Linacre 2017) a analýzu hlavních komponent pro určení unidimenzionality škály (Baker a kol., 2004, Linacre 2017, Smith 2002).

Výsledná motorická škála by měla být určena pro trenéry florbalu kategorie mladšího školního věku. Škála je zkonstruována za účelem zjišťování úrovně florbalové techniky, ovládnutí míčku pomocí florbalové hole, která determinuje možnosti hráče při hře. Informace o úrovni dovedností u jednotlivých hráčů by měla trenérům poskytnout zpětnou vazbu v tréninkovém procesu začínajících florbalistů.

Postup při konstrukci škály by měl být také přínosný pro další možnosti tvorby motorických škál v ostatních sportovních specializacích a kinantropologii.

2. TEORETICKÝ ZÁKLAD ŘEŠENÍ

Cílem hráčů ve florbalu je vstřelit více branek než soupeř, to znamená ovládat míček na takové úrovni, aby byl dopraven do branky soupeře. Manipulace s míčkem, navzdory bránění soupeře, a vstřelení míčku do branky, kterou brání brankář, je bez odpovídajících technických dovedností, stejně jako je tomu v ostatních kolektivních hrách, téměř nemožná.

Nezanedbatelným faktorem, který souvisí se specifickými sportovními dovednostmi a jejich technickým provedením, je faktor techniky. Jako u ostatních sportovních her je florbalový výkon založen v první řadě na vysoké úrovni zvládnutých dovedností. Dovednosti jsou z hlediska správného provedení pro výkon rozhodující. Kondiční schopnosti mohou úroveň dovedností podporovat nebo omezovat, nikoliv suplovat (Velenský, 2008).

Technikou se rozumí účelný způsob řešení pohybového úkonu, který je v souladu s možnostmi jedince, s biomechanickými zákonitostmi pohybu a uskutečňuje se na základě neurofyziologických mechanismů řízení pohybu. Využívají se přitom i další předpoklady sportovce, především kondiční, somatické i psychické. Sportovní dovednosti jsou základem pro specifické jednání v soutěžích, jejich technika čili způsob provedení, zásoba sportovních dovedností, stabilita i proměnlivost, při které hráč reaguje na vnější podmínky tak, aby úkol byl co nejlépe splněn, jsou významnými specifickými faktory struktury sportovního výkonu (Dovalil a kol., 2005).

Úroveň technických dovedností ve florbalu ovlivňuje dovednost ovládnutí florbalové hole a míčku (Kysel 2010, Skružný 2005, Martínková 2009). Mezi nezbytné technické dovednosti patří ale i způsob pohybu hráče, tj. technika běhu, technika zastavení a startu do různých směrů pohybu. Jednotlivé pohyby hráče s míčkem nebo za účelem získání míčku, používané při hře, nazýváme herními činnostmi jednotlivce a jsou základním stavebním prvkem herních kombinací a systémů (Kysel 2010, Skružný 2005, Martínková 2009, Zlatník 2004).

Cílem předkládané práce je konstrukce diagnostického nástroje, který bude zjišťovat úroveň technické dovednosti ovládnutí florbalové hole a míčku u hráčů florbalu. Zjišťování úrovně dovednosti ovládnutí florbalové hole a míčku je obtížně kvantifikovatelný rys. Po prostudování dostupných testů a postupů standardizace v kinantropologii jsme se rozhodli pro odhad úrovně florbalových dovedností použít posuzovací škálu. Po prostudování možností, které nabízí kinantropologie, jsme se rozhodli pro aplikaci škály Guttmanova typu. Jedná se o

sestavu položek řazených takovým způsobem, aby narůstala úroveň latentního rysu potřebná k jejich splnění. „To znamená, že testovaná osoba by měla splnit blok položek od počátku škály až po kritický bod, který značí maximální možnou úroveň latentního rysu u testované osoby“ (Čepička, 2005, 10). Jedná se o škálu kumulativní, položka nacházející se v kritickém bodě zaručuje úspěšné splnění všech předchozích položek. O úrovni latentního rysu u testované osoby vypovídá tedy poslední splněná položka. Úroveň dovedností na jednotlivých položkách je hodnocena dichotomicky, tj. ve smyslu 1 – 0, splnil – nesplnil nebo správně – chybně.

Základním předpokladem Guttmanovy škály je unidimenzionalita, která znamená, že všechny položky škály diagnostikující shodný latentní rys. „Unidimenzionalita položek je tak limitujícím faktorem pro tvorbu škály Guttmanova typu“ (Čepička, 2005, 10). Latentní rys, neboli latentní proměnná, je přímo nepozorovatelná a můžeme jí přisoudit věcný smysl v podobě obecné vlastnosti jako je rys osobnosti, pohybová schopnost (Blahuš, 1996) nebo v případě naší práce pohybová dovednost. „Cílem modelování s latentními proměnnými je při zadaných hodnotách manifestních, přímo pozorovatelných, proměnných matematicky zkonstruovat neznámou soustavu latentních proměnných a jejich vhodnou funkci, která by dostatečně dobře aproximovala buď přímo samotné manifestní proměnné a jejich hodnoty anebo eventuálně hustotu pravděpodobností manifestních proměnných, tedy (nepřesně řečeno) pravděpodobnost výskytu jejich hodnot“ (Blahuš, 1996, 183). K modelování pomocí latentních proměnných, kterých se používá při analýze testových výsledků binárního a obecně kategoriálního typu, se používá teorie latentních rysů neboli teorie položkových odpovědí. Teorie vychází z toho, že testové výsledky jsou indikátory latentních proměnných a zajímají nás pravděpodobnosti správného zodpovězení položky v závislosti na schopnostech jedince (Hendl, 2015).

Pomocí Raschova modelu se budeme snažit o vysvětlení vztahu mezi teoretickou vlastností reprezentovanou latentní proměnnou a empirickou vlastností reprezentovanou manifestní proměnnou. Latentní proměnnou je latentní rys florbalová dovednost ovládnutí florbalové hole a míčku u hráčů florbalu. Manifestní proměnnou je odpověď na dichotomické položce (1,0; tj. 1 odpovídá splnění motorického úkolu, 0 znamená nesplnění motorického úkolu).

Guttmanova škála je teoretický a matematický ideál. Jedná se o ordinální škálu, která neposkytuje žádné informace o intervalech mezi položkami nebo o intervalech mezi testovanými osobami. Z tohoto důvodu jsou měření vystavena chybám z překrývajícími se

proměnných, zvláště když se jedná o jemnější stupnice. Jako prostředek pro odhad intervalových stupnic, určení obtížnosti jednotlivých položek, se využívají právě modely teorie položkových odpovědí, tj. IRT modely. (Massof, 2011) „Pomocí teorie položkových odpovědí lze sofistikovaným způsobem také posuzovat vhodnost položek pro konkrétní položkovou sestavu, navrhnout jejich úpravy, doplnění či vypuštění ze sestavy“ (Čepička, 2005). Raschovy modely teorie položkových odpovědí jsou charakteristické právě nezávislostí jednotlivých položek, dále pak nezávislostí testovaných osob, společnou nezávislostí položek a testovaných osob a lokální nezávislostí (Baker a kol., 2004). Základním modelem je jednoparametrický Raschův model, který poskytuje přímo vyjádřené srovnání hodnot úrovně latentního rysu a obtížnosti položky.

3. CÍL PRÁCE A ÚKOLY PRÁCE

Předložená práce vychází z potřeby objektivního hodnocení florbalových dovedností ve florbalu. Naší snahou je vytvořit diagnostický nástroj, který bude použitelný pro trenéry florbalu v běžné sportovní praxi. Mezi hlavní požadavky úspěšného řešení řadíme praktičnost sestrojeného diagnostického nástroje. Výsledkem práce by také měl být vytvořený postup konstrukce motorické škály využitelný v kinantropologii a v ostatních sportovních specializacích.

Cíl práce

Cílem práce je návrh standardizovaného diagnostického nástroje, posuzovací škály, určené pro diagnostiku florbalových dovedností hráčů florbalu mladšího školního věku.

Úkoly práce

Úkoly práce jsme rozdělili do dvou fází:

1. V první fázi projektu přistoupíme ke konstrukci diagnostického nástroje.

Dílčí úkoly:

- vybrat nebo vytvořit položky určené k testování florbalových dovedností,
- sestavit tým florbalových expertů a provést expertní analýzu,
- identifikovat položky s nejvyšší obsahovou validitou,
- ověřit položky při pilotním testování,
- provést hlavní sběr dat a analýzu dat,
- vybrat položky do výsledné škály.

2. Ve druhé fázi projektu provedeme vědeckou standardizaci diagnostického nástroje.

Dílčí úkoly:

- posoudit validitu diagnostického nástroje,
- vypočítat reliabilitu diagnostického nástroje,
- zhodnotit diagnostický nástroj z hlediska objektivity.

4. METODOLOGIE VÝZKUMU

4.1 Výzkumný soubor

Výzkumný soubor byl složen z 212 hráčů vybraných florbalových družstev. Výzkumu se zúčastnilo 42 hráčů kategorie přípravky (6 – 8 let), 141 hráčů kategorie elévové (9 – 10 let) a 29 mladších žáků (11 – 12 let) ze Středočeského kraje. Výzkumný soubor tvořilo 197 chlapců a 15 děvčat, které s chlapci ve výše uvedených kategoriích trénovaly.

Pilotního testování, které ověřovalo proceduru testování a vhodnost vybraných položek, se zúčastnilo 24 hráčů ve věku 6 – 12 let.

Pro vytvoření motorické škály je žádoucí, aby kalibrační soubor nebyl z hlediska diagnostikovaného jevu homogenní a v testovaném souboru byl obsažen co možná nejširší úsek kontinua kvantifikujícího zkoumaný předpoklad (Čepička, 2005). Kategorie mladšího školního věku by tento požadavek, z hlediska výše uvedeného nestejného motorického vývoje v období dětství a prepubescence (Perič, 2012), měla splňovat. Do souboru byli také zařazeni hráči s nestejnou úrovní osvojených florbalových dovedností. Soubor byl složen z hráčů soutěžních a nesoutěžních oddílů. Nesoutěžní oddíly jsou kroužky florbalu, ve kterých se hráči nezúčastňují ligových utkání, jedná se o rekreační formu florbalu. V soutěžních oddílech se hráči účastní turnajů a jedná se o výkonnostní sport.

Přehled probandů z hlediska dovednostní úrovně:

- 36 hráčů, kteří s florbalovým tréninkem začínají, zúčastňují se tréninků méně než 1 rok;
- 48 hráčů, kteří trénují v nesoutěžním družstvu 1–2 roky, 1x nebo 2x týdně;
- 39 hráčů, kteří trénují v nesoutěžním družstvu 3–4 roků, 1x nebo 2x týdně;
- 38 hráčů, kteří trénují v soutěžním družstvu 1–2 roky, 3x týdně;
- 36 hráčů, kteří trénují v soutěžním družstvu 3–4 roků, 3x týdně;
- 15 hráčů, kteří trénují v soutěžním družstvu 5 a více roků, 3x týdně.

4.2 Výzkumné metody

Pro účely práce jsme vycházeli z postupů pro standardizaci motorické škály navržené Čepičkou (2002, 2005) a Štochlem a Musálkem (2009).

K výzkumu jsme také použili následující metody statistické analýzy určené k posouzení vlastností a unidimenzionality škály: expertní analýzu pro posouzení obsahové validity podle Lawshe (1975), ztrátové fit funkce pro určení shody modelu s daty (Blahuš 1996, Linacre 2017, Smith 1998), Fleissovo kappa pro zjištění shody posuzovatelů (Fleiss, 1971), koeficient KR – 20 pro určení reliability škály (Cortina, 1993, Cronbach 1951, Linacre 2017) a analýzu hlavních komponent pro určení unidimenzionality škály (Baker a kol. 2004, Linacre 2017, Smith 2002).

Pro účely disertační práce jsme analýzu dat provedli pomocí programu Winsteps, verze 4.0. z roku 2017, který je přímo určen pro analýzu binárních dat pomocí Raschova modelu. Rozsah dat, s kterými může program pracovat, je 60 000 položek a 10 000 000 osob. K analýze lze využít nejen dichotomického Raschova modelu ale i například Partial Credit, Rating Scale nebo Generalized Partial Credit modelů. Pro výpočet Fleissova kappa jsme použili počítačový software Online Kappa Calculator (Randolph, 2008).

4.2.1 Postup při konstrukci motorické škály

Proces definování podmínek použití konkrétního modelu, které určují prostor pro interpretaci testových výsledků, se obecně nazývá kalibrací. Jestliže se rozhodneme pro použití určitého modelu pro interpretaci vztahu mezi odpovědí na položku a úrovní latentního rysu, budeme muset odhadnout hodnotu parametrů položkové funkce a shodu vybraného modelu s daty (Čepička, 2002). Čepička navrhuje následující postup:

- a) návržení metodiky, hypotéz a designu výzkumu včetně výběru modelu;
- b) sběr dat (pro jednoparametrový model se doporučuje minimální rozsah výběru $N=100$);

- c) provedení odhadu parametrů na základě sebraných dat (v případě Raschova modelu pouze parametru obtížnosti);
- d) posouzení vhodnosti modelu ve smyslu jeho shody s daty.

Tento obecný postup jsme rozšířili o jednotlivé kroky navržené Štochlem a Musálkem (2009) a postup jsme doplnili o posouzení unidimenzionality škály jako nutné podmínky k sestavení škály Guttmanova typu. Jednotlivé kroky jsou zaznamenány ve *schématu č. 2*.

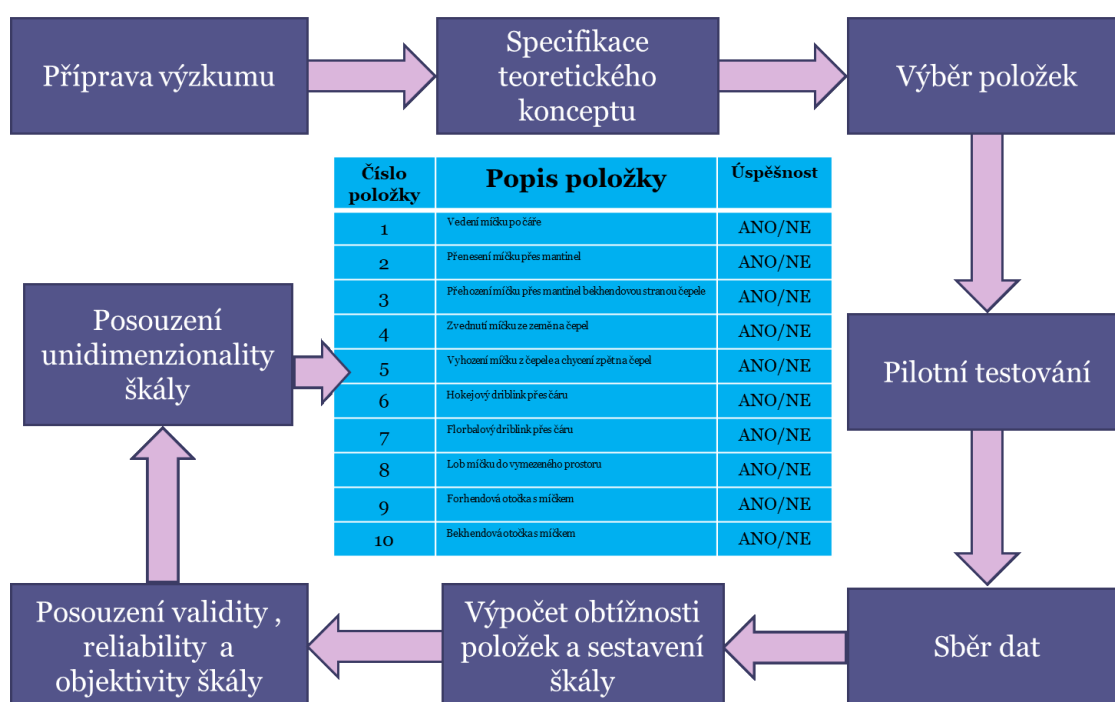


Schéma č. 2 Postup při standardizaci škály pro posouzení úrovně florbalové dovednosti

4.2.2 Posouzení validity, reliability, objektivit a unidimenzionalit škály

V teorii položkových odpovědí je validita posuzována společně s dimenzionalitou diagnostického nástroje nebo jednotlivých položek, výsledky odpovědi na položku jsou tedy odrazem jednoho teoretického konceptu u testované osoby. Míra shody tohoto předpokladu s reálnými daty je posuzována fit statistikami. Při použití Raschova modelu pro odhad

parametrů položek je tak posouzení teoretické validity záležitostí ověření vhodnosti použití modelu pro analyzovaný soubor dat, tj. Raschův model je nástrojem pro posouzení teoretické validity (Čepička, 2005). Jestliže tedy získáme informace o shodě modelu s daty, získáme také informaci o teoretické validitě škály.

Při zkoumání infit a outfit statistik u jednotlivých položek doporučuje Linacre (2017) postupovat podle pravidel:

1. zkoumejte nejprve hodnoty outfit před infit,
2. hodnoty MNSQ před ZSTD,
3. vysoké hodnoty před nízkými nebo zápornými hodnotami.

Hodnota infit (respektive outfit) MNSQ znamená střední kvadratické reziduum vypočítané z hodnoty χ^2 dělené stupni volnosti. Hodnota infit (respektive outfit) ZSTD značí standardizovanou hodnotu rezidua, tj. pravděpodobnost MNSQ vyjádřenou v Z bodech.

Linacre (2017) upozorňuje, že vysoké hodnoty MNSQ nebo kladné hodnoty ZSTD jsou mnohem větší hrozbou vůči validitě než nižší respektive záporné hodnoty. Pro ideální shodu modelu s daty se očekává hodnota infit a outfit statistik rovná 1. Čepička (2003, 2005) a Zhu (1996) považovali hodnoty v intervalu (0,7;1,3) za vyhovující, Linacre (2017) považuje za vyhovující hodnoty z intervalu (0,5;1,5), Wilson (2005) hodnoty (0,7;1,33). Pokud jsou hodnoty menší než 1, nemusí se podle Linacreho (2017) jednat o nevyhovující model, ale spíše o zkreslená data. Hodnoty ZSTD by se v ideálním případě měly rovnat 0 a záporné hodnoty odkazují spíše na přílišnou předvídatelnost dat.

Raschova analýza nám poskytuje fit statistiky k testování předpokladů měření, tj. probandi s vysokou úrovní dovednosti by měli být úspěšní na téměř všech jednodušších položkách. Jakmile jsou identifikovány osoby a položky, které neodpovídají modelu, mohou být změněny nebo odstraněny (Linacre 2017, McCreary a kol. 2014).

Linacre (2017) doporučuje nejprve zkoumat outfit statistiky před infit statistikami, MNSQ statistiky před ZSTD statistikami a kladné hodnoty před zápornými nebo negativními hodnotami. Dále uvádí, že důvodem pro hodnocení nejprve kladných hodnot MNSQ před zápornými je skutečnost, že MNSQ fit statistiky se zprůměrují vždy na hodnotu 1, tak pokud budeme akceptovat položky s vysokými hodnotami (tj. nízkou diskriminativitou), pak zároveň musíme akceptovat položky s nízkou hodnotou (tj. vysokou diskriminativitou). A zatímco

vysoké outfit MNSQ hodnoty jsou většinou jednoduše diagnostikovatelnými a přepracovatelnými, tak vysoké hodnoty infit MNSQ jsou špatně diagnostikovanými a jsou mnohem větší hrozbou pro validitu měřícího nástroje. Vysoké hodnoty znamenají, že položky jsou zavádějící pro skupinu osob, pro kterou jsou určeny. Hodnoty MNSQ nám celkově ukazují velikost zkreslení měřícího nástroje, hodnoty menší než 1,0 nám poukazují na předvídatelné hodnoty a hodnoty vyšší než 1,0 indikují u položek nepředvídatelnost.

Linacre (2017) interpretuje hodnoty parametrů MNSQ fit statistik následujícím způsobem:

Tabulka č. 2 Hodnoty MNSQ fit statistik (Linacre, 2017)

Hodnota MNSQ fit statistik	Interpretace podle Linacre (2017)
> 2	Vyšší hodnoty zkreslují nebo zbavují hodnoty měřící přístroj.
1,5 – 2,0	Hodnota parametru nevhodná pro konstrukci měřícího nástroje, ale nedegraduje měřící nástroj.
0,5 – 1,5	Vhodná hodnota pro měření.
< 0,5	Hodnota parametru méně vhodná pro konstrukci měřícího nástroje, ale nedegraduje měřící nástroj. Může vytvářet zavádějící hodnoty dobré reliability a separace.

Linacre (2017) dále uvádí, že hodnoty středního kvadratického rezidua naznačují odchylku od unidimenzionality v datech, nikoliv v měření a výsledkem jsou zkreslená data. Dále uvádí, že vysoké outfit hodnoty MNSQ mohou být výsledkem několika náhodných odpovědí probandů s nízkou úrovní zkoumané dovednosti. Doporučuje proto, stejně jako (McCreary a kol., 2014), vyřadit tyto probandy ze zkoumaného souboru.

ZSTD, standardizovaná hodnota rezidua, jsou t-testy hypotézy: „Odpovídají přesně data modelu?“ Očekávané hodnoty by měly nabývat hodnoty 0,0. Nižší hodnoty indikují velkou předvídatelnost dat, vyšší hodnoty naopak nedostatek předvídatelnosti. Pokud jsou hodnoty MNSQ přijatelné, potom se hodnoty ZSTD nemusí brát v úvahu (Linacre, 2017) a u některých studií nejsou uváděny (např. McCreary a kol., 2014).

Předpoklad unidimenzionality u Raschova modelu, tj. základní předpoklad Guttmanovy škály, je možné testovat různými způsoby. Výsledky lze pak považovat za řadu ukazatelů podporujících nebo naopak vyvracejících předpoklad unidimenzionality. V naší

práci budeme postupovat ve dvou krocích, které byly použity například u Raschovy analýzy ve studiích Donovan a kol. (2006), McCreary a kol. (2014) nebo Wu a kol. (2016). Nejprve využijeme fit statistik a předpokladu, že pokud se shodují data s modelem, pak by měl být předpoklad unidimenzionality splněn (Čepička 2005, Linacre 2017). V druhém kroku použijeme analýzu hlavních komponent standardizovaných reziduí k určení případné jiné dimenze u položek (Smith, 2002).

U analýzy hlavních komponent jsme vycházeli z doporučení Linacreho (2017), Embretsona a Reise (Embretsona a Reise 2000 v McCreary a kol., 2014, 6) a Raicheho (Raiche 2005 v Basílio a kol., 2016, 7) a zvolili jsme tři kritéria pro hodnocení unidimenzionality:

- Raschova dimenze, první hlavní komponenta, bude vysvětlovat nejméně 50% celkového rozptylu v datech;
- druhá největší komponenta, tj. první kontrast v reziduích, nebude větší než 5 % ze zbylého rozptylu nebo hodnota *eigenvalue*, tzv. vlastního čísla, bude < 2 ;
- minimální poměr 3:1 bude mezi rozptylem vysvětleným položkami a rozptylem vysvětleným prvním kontrastem, druhou největší komponentou.

Reliabilitu testu budeme hodnotit pomocí tzv. klasické teorie testů, respektive koeficientu KR – 20. Hodnoty koeficientu budeme interpretovat podle Tavakola a Dennicka (2011) viz *tabulka č. 3*, která popisuje dosažené hodnoty Crombachova alfa, obecné verze Kuder – Richardsonova koeficientu ekvivalence (KR – 20).

Tabulka č. 3 Hodnoty Crombachova alfa (Tavakol a Dennick, 2011)

Hodnota koeficientu α	Interpretace podle Tavakola a Dennicka (2011)
$\alpha \geq 0,9$	výborná
$0,9 > \alpha \geq 0,8$	dobrá
$0,8 > \alpha \geq 0,7$	přijatelná
$0,7 > \alpha \geq 0,6$	diskutabilní
$0,6 > \alpha \geq 0,5$	slabá
$\alpha > 0,5$	nepřijatelná

Linacre (2017) uvádí, že reliabilita testu interpretována pomocí klasické teorie testů (koeficientu KR – 20) dosahuje o něco vyšších hodnot než Raschova reliabilita.

Pokud Raschova reliabilita osob dosahuje hodnot $< 0,8$ a separační koeficient osob dosahuje hodnot < 2 , potom při dostatečně velkém vzorku osob nám výsledky naznačují, že měřicí nástroj není dostatečně citlivý, aby rozlišoval mezi osobami skórujícími na vysokých a nízkých položkách. Linacre (2017) pak doporučuje rozšířit nástroj o více položek. Uvádí, že vysoká reliabilita osob závisí především na vzorku s dostatečně rozsáhlým kontinuem kvantifikující zkoumaný předpoklad testovaného souboru, dostatečné délce testu a správným skórováním osob na odpovídajících položkách.

Pokud Raschova reliabilita položek dosahuje hodnot $< 0,9$ a separační koeficient položek dosahuje hodnot < 3 , potom nám výsledky naznačují, že vzorek osob není dostatečně velký, aby potvrdil posloupnost obtížnosti položek. Vysoká reliabilita položek tedy závisí především na dostatečném počtu různě obtížných položek a také na velikosti testovaného souboru.

Ve výsledcích se uvádějí také hodnoty „reálné“ a „modelové“ reliability osob a položek. Modelová reliabilita je reliabilita vypočtená pro případ, že by struktura dat ideálně vyhovovala Raschově modelu a tvoří horní hranici intervalu, ve kterém leží skutečná hodnota reliability. Naopak reálná reliabilita osob a položek je vypočtená z dat sebraných testováním reálného souboru testovaných osob a tvoří pak spodní hranici tohoto intervalu. (Čepička, 2005)

Pro hodnocení objektivitu, ke kterému jsme použili Fleissovo kappa, jsme zvolili interpretaci podle Landise a Kocha (1977). Hodnoty jsou zaznamenány v *tabulce č. 4*.

Tabulka č. 4 Hodnoty Fleissova kappa (Landis a Koch, 1977)

Hodnota Fleissovo κ	Interpretace podle Landise a Kocha (1977)
<0	slabá shoda
0,0-0,20	nepatrná shoda
0,21-0,40	uspokojivá shoda
0,41-0,60	průměrná shoda
0,61-0,80	dobrá shoda
0,81-1,0	téměř dokonalá shoda

Objektivitu budeme hodnotit pro každou položku výsledné škály zvlášť, tj. „*item-by-item inter-rater agreement analysis*“ (Turner–Stokes a kol., 2014). Pro celou škálu uvedeme průměrnou hodnotu Fleissova kappa vypočítanou z hodnot u jednotlivých položek, a také pozorovanou celkovou procentuální shodu examinátorů u jednotlivých položek (Welch, 2012) a průměrnou hodnotu u celé škály.

5. VÝSLEDKY

5.1 Výsledky expertní analýzy

Výsledky expertní analýzy, které se zúčastnilo 5 expertů jsme zaznamenali do *tabulky* č. 5. U každé z 30 položek jsme vypočítali koeficient podle Lawshe (1975), CVR, a položky seřadili podle hodnoty koeficientu od nejvyššího po nejnižší.

Popis položek vychází z obvyklé terminologie specializace florbalu:

1. Vedení míčku po rovné čáře.
2. Vedení míčku se změnou směru (po obvodu čtverce).
3. Driblink míčku přes širokou čáru – hokejový driblink.
4. Driblink míčku přes širokou čáru – florbalový driblink.
5. Vedení míčku po trajektorii osmičky kolem dvou kuželů.
6. Vedení míčku po trajektorii osmičky mezi nohama.
7. Otočka s míčkem na forhendovou stranu.
8. Otočka s míčkem na bekhendovou stranu.
9. Slalom č. 1 – vedení míčku mezi kužely v časovém limitu (15 kuželů je umístěno na čáře za sebou ve vzdálenosti 30 cm).
10. Slalom č. 2 – vedení míčku mezi kužely v časovém limitu (6 kuželů je umístěno na dvou rovnoběžných čarách vzdálených 50 cm, kužely jsou umístěné na přeskáčku na levé a pravé čáře ve vzdálenosti 50 cm od sebe).
11. Slalom č. 3 – vedení míčku mezi kužely v časovém limitu (tři kužely jsou umístěné vždy za sebou ve vzdálenosti 30 cm na jedné čáře a na ně navazují tři kužely na druhé čáře a opět tři kužely na první čáře a tři kužely na druhé čáře, čáry jsou rovnoběžné a vzdálené 50 cm, hráč vede míček nejprve mezi kužely po levém boku, potom vede míček mezi kužely po pravém boku a to se opakuje celkem 2x).
12. Člunkový běh na čas mezi dvěma čarami.
13. Člunkový běh na čas mezi dvěma kužely.
14. Zvednutí míčku na čepel florbalové hole.
15. Otočka s míčkem ve vzduchu (míček je položený na čepeli).
16. Vyhození míčku z čepele do vzduchu a chycení zpět na čepel.

17. Driblink míčku ve vzduchu z forhendové strany čepele (počet pokusů bez dopadu na zem).
18. Driblink míčku ve vzduchu z bekhendové strany čepele (počet pokusů bez dopadu na zem).
19. Přenášení míčku přes mantinel (hráč stojí bokem k mantinelu u jeho jednoho konce a přenáší míček z jedné strany na druhou, mantinel nepřekračuje).
20. Přehození mantinelu forhendovou stranou čepele.
21. Přehození mantinelu bekhendovou stranou čepele.
22. Přihrání míčku mezi dva kužely.
23. Přihrání míčku z běhu mezi dva kužely.
24. Přihrávání několika míčků z běhu a mezi dva kužele během časového limitu.
25. Střelba na branku z místa (v brance visí obruč, do které se hráč musí trefit).
26. Střelba na branku z běhu (v brance visí obruč, do které se hráč musí trefit).
27. Střelba na branku několika míčků z běhu a během časového limitu (v brance visí obruč, do které se hráč musí trefit).
28. Střelba na branku do označených sektorů.
29. Lob na vzdálený cíl (lob míčku do vymezeného prostoru)).
30. Lob na vzdálenost.

Tabulka č. 5 Koeficient CVR podle Lawshe (1975)

Číslo položky	CVR
Položka č.16	1
Položka č.22	1
Položka č.25	1
Položka č.3	0,6
Položka č.4	0,6
Položka č.6	0,6
Položka č.8	0,6
Položka č.9	0,6
Položka č.10	0,6
Položka č.11	0,6
Položka č.12	0,6
Položka č.28	0,6
Položka č.29	0,6
Položka č.1	0,2
Položka č.5	0,2
Položka č.7	0,2
Položka č.14	0,2
Položka č.17	0,2
Položka č.18	0,2
Položka č.19	0,2
Položka č.21	0,2
Položka č.23	0,2
Položka č.24	0,2
Položka č.26	0,2
Položka č.27	0,2
Položka č.13	-0,2
Položka č.20	-0,2
Položka č.2	-0,6
Položka č.15	-1
Položka č.30	-1

Položky se třemi nejvyššími koeficienty 1; 0,6 a 0,2 jsme zařadili do sestavy položek určené k testování. Vyřadili jsme položky č. 2, 13, 15 20 a 30 se zápornými koeficienty -1; -0,6 a -0,2 a také položky č. 5, 12, 17, 18, 24, 27 a 28 s vyššími koeficienty, tak abychom počet položek snížili na nižší počet, který jsme z časových důvodů schopni na probandech

otestovat při pilotním testování. Položky s vyššími koeficienty nebyly zařazeny do sestavy položek určené k testování na základě diskuse s experty.

Pro pilotní testování bylo na základě expertní analýzy vybráno 18 následujících položek, které jsme očíslovali v následujícím pořadí:

Položka č. 1 – vedení míčku po rovné čáře

Položka č. 2 – driblink míčku přes širokou čáru – hokejový driblink

Položka č. 3 – driblink míčku přes širokou čáru – florbalový driblink

Položka č. 4 – vedení míčku po trajektorii osmičky mezi nohama

Položka č. 5 – otočka s míčkem na forhendovou stranu

Položka č. 6 – otočka s míčkem na bekhendovou stranu

Položka č. 7 – vedení míčku mezi kužely v časovém limitu (slalom č. 1)

Položka č. 8 – vedení míčku mezi kužely v časovém limitu (slalom č. 2)

Položka č. 9 – vedení míčku mezi kužely v časovém limitu (slalom č. 3)

Položka č. 10 – zvednutí míčku na čepel florbalové hole

Položka č. 11 – vyhození míčku z čepele do vzduchu a chycení zpět na čepel

Položka č. 12 – přenášení míčku přes mantinel

Položka č. 13 – přehození mantinelu bekhendovou stranou čepele

Položka č. 14 – přihrání míčku mezi dva kužely

Položka č. 15 – přihrání míčku z běhu mezi dva kužely

Položka č. 16 – střelba na branku z místa (do obruče)

Položka č. 17 – střelba na branku z běhu (do obruče)

Položka č. 18 – lob na vzdálený cíl (lob míčku do vymezeného prostoru)

5.2 Výsledky pilotního testování

U 25 probandů mladšího školního věku jsme ověřovali vhodnost 18 výše vybraných položek z hlediska adekvátní obtížnosti pro mladší školní věk a z hlediska vhodnosti k testování. Z hlediska obtížnosti jsme ohodnotili všechny položky jako vhodné pro tuto věkovou kategorii. Položky č. 3 a č. 4 jsme z hlediska možnosti testování vyhodnotili jako nevyhovující a pro hlavní testování jsme využili zbylých 16 položek, které jsme přečíslovali a zjednodušili jejich název, respektive popis (*tabulka č. 6*).

Tabulka č. 6 Položky určené k hlavnímu testování

Číslo položky	Název položky
Položka č.1	forhendová otočka
Položka č.2	bekhendová otočka
Položka č.3	osmička
Položka č.4	zvednutí míčku na čepel
Položka č.5	vyhození míčku z čepele a chycení zpět
Položka č.6	přenesení míčku přes mantinel
Položka č.7	přehození míčku přes mantinel bekhendovou stranou čepele
Položka č.8	vedení míčku po čáře
Položka č.9	lob míčku do vymezeného prostoru
Položka č.10	příhrávka z místa na cíl
Položka č.11	příhrávka z pohybu na cíl
Položka č.12	střelba z místa na cíl
Položka č.13	střelba z pohybu na cíl
Položka č.14	slalom s míčkem č. 1
Položka č.15	slalom s míčkem č. 2
Položka č.16	slalom s míčkem č. 3

5.3 Sestavení škály

Sestavení škály jsme prováděli ve třech krocích, nejprve jsme analyzovali fit statistiky u jednotlivých položek, potom jsme vytvořili grafickou podobu škály podle parametru obtížnosti jednotlivých položek a nakonec zhodnotili škálové hodnoty.

5.3.1 1. fáze – vyřazení nevyhovujících položek

Tabulka č. 7 nám ukazuje výsledné hodnoty fit statistik pro výzkumný soubor 212 hráčů hodnocených na 16 testovacích položkách.

Tabulka č. 7 Výsledek Raschovy analýzy v souboru 212 probandů testovaných na 16 položkách

POLOŽKA	OBTÍŽNOST POLOŽKY	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
13	4.34	0.80	-0.9	0.25	-1.6
11	3.68	0.94	-0.3	1.12	0.4
12	3.38	0.80	-1.3	0.31	-1.6
8	2.80	1.49	3.2	1.56	1.1
10	2.13	1.27	2.1	2.15	2.1
9	1.59	1.00	0.0	0.82	-0.3
14	1.51	0.77	-2.2	0.48	-1.5
16	0.42	1.11	1.1	0.93	-0.1
5	-0.12	0.71	-3.3	0.48	-2.3
4	-0.88	0.82	-1.9	0.55	-1.7
15	-1.77	1.20	1.8	1.20	0.6
3	-2.39	1.00	0.1	1.11	0.4
7	-2.78	0.92	-0.6	0.58	-0.9
6	-3.11	1.00	0.1	0.82	-0.2
2	-3.22	1.05	0.4	2.42	2.1
1	-5.59	0.92	-0.2	1.03	0.3

V hodnocení jsme postupovali podle Linacra (2017), který doporučuje zkoumat nejprve hodnoty outfit před infit, hodnoty MNSQ před ZSTD a nakonec vysoké hodnoty před nízkými nebo zápornými hodnotami. Za vyhovující hodnoty považuje interval (0,5;1,5).

Hodnoty outfit MNSQ, které nevyhovují předepsanému intervalu, jsme našli u položek č. 13, č. 12, č. 8, č. 10, č. 14, č. 5 a č. 2. Položky č. 14 a č. 5 mají hodnotu outfit MNSQ 0,48. Jedná se o minimální vybočení z intervalu (hodnoty $< 0,5$) a navíc hodnoty $< 0,5$ měřící nástroj nedegradují, jsou pouze méně vhodné pro jeho konstrukci (viz kapitola 6.3.7.). Z tohoto důvodu jsme se rozhodli položky pro další zkoumání zachovat, hrozbou pro unidimenzionalitu jsou hodnoty $> 1,5$ a nikoliv nízké hodnoty. Testovou položku č. 13, č. 12 a č. 2 s hodnotou outfit MNSQ mimo doporučený interval (0,5;1,5) jsme se rozhodli ze škály vyřadit, protože i hodnoty parametrů obtížnosti položek jsou příliš nízké nebo příliš vysoké (mimo vhodný interval od -3 do 3 logitů), aby mohly být zařazeny do výsledné škály.

S položkou č. 10 a č. 8 jsme se ale rozhodli dále pracovat, protože hodnota její obtížnosti vhodně doplňuje vzdálenosti mezi jednotlivými položkami ve škále. Linacre (2017) doporučuje pro snížení hodnoty outfit MNSQ se pokusit identifikovat osoby s nízkou úrovní latentního rysu, kteří byli náhodně úspěšní na obtížných položkách, a tyto osoby vyřadit. Vysoké hodnoty tak nemusejí znamenat hrozbu pro unidimezionalitu, ale jsou výsledkem několika náhodných odpovědí.

Pro prozkoumání testových výsledků probandů jsme ze souboru vyřadili probandy č. 4, č. 68, č. 79 a č. 90, kteří náhodně skórovali na obtížných položkách č. 10 (probandi č. 4, č. 68 a č. 90) a č. 8 (proband č. 79). Výsledné hodnoty fit statistiky jsou zaznamenány v *tabulce* č. 8. Hodnota outfit MNSQ na položce č. 10 klesla na 1,04 a na položce č. 8 na 1,14, a tím splňují doporučený interval. Neplánovaně se mírně navýšily i hodnoty na položkách č. 14 a č. 5 na přijatelnější hodnoty 0,49.

Tabulka č. 8 Výsledek Raschovy analýzy v souboru 208 probandů testovaných na 16 položkách

POLOŽKA	OBTÍŽNOST POLOŽKY	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
13	4.39	0.80	-0.9	0.25	-1.5
11	3.72	0.95	-0.2	1.19	0.5
12	3.41	0.80	-1.3	0.31	-1.6
8	2.90	1.48	3.1	1.14	0.4
10	2.30	1.18	1.4	1.04	0.2
9	1.61	1.01	0.2	0.86	-0.2
14	1.52	0.78	-2.1	0.49	-1.4
16	0.41	1.14	1.3	0.97	0.0
5	-0.15	0.73	-3.0	0.49	-2.1
4	-0.93	0.85	-1.6	0.57	-1.6
15	-1.87	1.25	2.1	1.26	0.7
3	-2.38	1.01	0.1	1.13	0.4
7	-2.85	0.92	-0.6	0.58	-0.8
2	-3.20	1.07	0.6	2.49	2.2
6	-3.26	1.00	0.0	0.87	-0.1
1	-5.61	0.93	-0.1	1.05	0.3

Z tabulky č. 8 je také zřejmé, že všechny položky splňují hodnoty infit MNSQ v doporučeném intervalu (0,5;1,5). Vysoké infit MNSQ hodnoty jsou největší hrozbou pro validitu škály a nejsou jednoduše napravitelnými jako u hodnot outfit MNSQ. Tento výsledek je vzhledem ke konstrukci škály velmi pozitivní.

Hodnoty outfit ZSTD nabývají často nízkých nebo záporných nevyhovujících hodnot (položka č. 13, č. 12, č. 8, č. 10, č. 9, č. 14, č. 16, č. 5, č. 4, č. 3, č. 7, č. 6, č. 1). Vycházíme ale z tvrzení Linacreho (2017), že pokud jsou hodnoty MNSQ přijatelné, potom se hodnoty ZSTD nemusí brát v úvahu.

Při hodnocení parametrů obtížnosti, jejichž hodnoty by se měly pohybovat v intervalu od -3 logitů do +3 logitů, jsme jako nevyhovující položky shledali položky č. 1, č. 6, č. 2, které jsou příliš snadné, a položky č. 12, č. 11, č. 13, které jsou naopak příliš obtížné (tabulka č. 8).

Z měřicího nástroje bylo tedy na základě zkoumání fit statistik a parametru obtížnosti vyřazeno šest následujících položek: č. 1, č. 2, č. 6, č. 11, č. 12, č. 13.

5.3.2 2. fáze – kvantifikace obtížnosti jednotlivých položek

V další fázi v konstrukci diagnostického nástroje jsme pokračovali s 10 položkami, které vyhovovaly výše uvedeným kritériím, a podle hodnoty parametru obtížnosti jsme sestavili škálu od nejjednodušších položek po nejobtížnější (*tabulka č. 9*).

Tabulka č. 9 Škála uspořádaná podle parametrů obtížnosti pro 10 položek

POLOŽKA	8	10	9	14	16	5	4	15	3	7
OBTÍŽNOST	2.90	2.31	1.62	1.53	0.41	-0.17	-1.00	-1.99	-2.54	-3.06

Z *tabulky č. 9* vyplývá, že dvě položky (č. 9 a č. 14) mají parametr obtížnosti obdobné hodnoty (1,62 a 1,53 logitů). Položky ve výsledné škále by měly být pravidelně rozloženy po celém kontinuu škály, a proto jsme museli jednu ze stejně obtížných položek vyřadit. U testu č. 9 se infit statistiky ideálně přibližují hodnotám 1 (u MNSQ statistik) a 0 (u ZSTD statistik), navíc se jedná o test „lob míčku do vymezeného prostoru“, zatímco test č. 14 nemá tak vhodné fit statistiky a jedná se o slalomovou dráhu, která je již v jiných modifikacích zastoupená u položky č. 15 a č. 16. Položku č. 14 jsme ze škály vyřadili.

Položka č. 7 mírně překračuje interval parametru obtížnosti, nachází se na hranici velké jednoduchosti položek. Pro další zkoumání jsme se ale rozhodli položku č. 7 ve škále zachovat.

Hodnocení rozložení škálových hodnot jsme provedli u 9 zbylých položek, bez vyřazené položky č. 14.

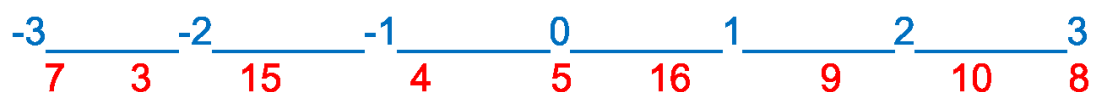
5.3.3 3. fáze – zhodnocení rozložení škálových hodnot

Nejprve jsme u 9 zbylých položek přepočítali hodnoty parametrů obtížnosti (*tabulka č. 10*) a vytvořili jsme grafickou podobu škály (*graf č. 2*), na které můžeme pozorovat distance škálových hodnot i s pokrytím kontinua diagnostikovaného rysu.

Tabulka č. 10 Škála uspořádaná podle parametrů obtížnosti pro 9 položek

POLOŽKA	8	10	9	16	5	4	15	3	7
OBTÍŽNOST	3.01	2.43	1.76	0.57	0.01	-0.81	-1.79	-2.34	-2.85

Graf č. 2 Posuzovací škála (modrá barva) s devíti položkami (červená barva) uspořádaná od nejjednodušší položky po nejobtížnější položku.



Výsledná škála optimálně pokrývá kontinuum diagnostikovaného rysu, o něco větší je vzdálenost mezi položkami č. 9 a č. 16 (1,19 logit) a mezi položkou č. 4 a č. 15 (0,98 logit). I přes tyto vzdálenosti se výsledná škála jeví vzhledem k rozložení položek a hodnotám parametrů obtížnosti jako kvalitní diagnostický nástroj latentního rysu.

V následující *tabulce č. 11* je znázorněna devítipoložková škála od nejjednodušších položek po nejobtížnější, kterou jsme podstoupily vědecké standardizaci:

Tabulka č. 11 Výsledná škála devíti položek určená k vědecké standardizaci

Číslo položky	Název položky
Položka č.7	přehození míčku přes mantinel bekhendovou stranou čepele
Položka č.3	osmička
Položka č.15	slalom s míčkem č. 2
Položka č.4	zvednutí míčku na čepel
Položka č.5	vyhození míčku z čepele a chycení zpět
Položka č.16	slalom s míčkem č. 3
Položka č.9	lob míčku do vymezeného prostoru
Položka č.10	příhrávka z místa na cíl
Položka č.8	vedení míčku po čáře

5.4 Standardizace škály – validita

Důležitým výsledkem při Raschově analýze výsledné škály jsou vzhledem k validitě škály hodnoty středního kvadratického rezidua (MNSQ), které všechny (infit i outfit hodnoty), dosahují podle Linacreho (2017) vhodných hodnot, tj. nevybočují z intervalu (0,5;1,5). Pokud jsou tyto hodnoty přijatelné, není nutné dále brát v úvahu hodnoty ZSTD (McCreary a kol. 2014, Linacre 2017).

Všechny hodnoty infit a outfit MNSQ odpovídají výše uvedenému intervalu (*tabulka č. 12*). Na základě těchto výsledků a jejich interpretace podle Linacreho (2017) tak můžeme tvrdit, že se model shoduje s daty a škála je validní. Shoda modelu s daty by nám také měla potvrdit předpoklad unidimenzionality.

Tabulka č. 12 Raschova analýza škály – fit statistiky

POLOŽKA	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
8	1.28	2.0	1.07	0.3
10	1.04	0.4	0.95	0.1
9	0.98	-0.2	0.93	-0.1
16	1.07	0.7	0.81	-0.8
5	0.73	-3.0	0.53	-2.7
4	0.89	-1.1	0.66	-1.4
15	1.17	1.4	1.13	0.4
3	1.10	0.8	1.04	0.3
7	0.93	-0.5	0.65	-0.5

Předpoklad unidimenzionality jsme se snažili potvrdit ve dvou krocích (Donovan a kol. 2006, McCreary a kol. 2014, Wu a kol. 2016). Prvním bylo výše zmíněné využití fit statistik pro doložení shody modelu s daty. V druhém kroku jsme použili analýzu hlavních komponent standardizovaných reziduí k určení případné jiné dimenze u položek. Výsledky jsou zaznamenány v *tabulce č. 13*.

Tabulka č. 13 Analýza hlavních komponent standardizovaných reziduí

	EIGENVALUE	Pozorovaný rozptyl	Očekávaný rozptyl
Celkový hrubý rozptyl vysvětlený měřeními	10.6188	54.1%	54.4%
Hrubý rozptyl vysvětlený osobami	5.1400	26.2%	26.3%
Hrubý rozptyl vysvětlený položkami	5.4788	27.9%	28.1%
Nevysvětlitelný rozptyl v 1.kontrastu	1.7225	8.8 %	19.1%
Nevysvětlitelný rozptyl v 2.kontrastu	1.4461	7.4 %	16.1%
Nevysvětlitelný rozptyl v 3.kontrastu	1.2083	6.2 %	13.4%
Nevysvětlitelný rozptyl v 4.kontrastu	1.1030	5.6 %	12.3%
Nevysvětlitelný rozptyl v 5.kontrastu	1.0289	5.2 %	11.4%

Při analýze jsme hodnotili tři kritéria (viz kapitola 6.4.7.):

1. *Raschova dimenze, první hlavní komponenta, bude vysvětlovat nejméně 50% celkového rozptylu v datech – pozorovaný celkový hrubý rozptyl vysvětlený měřeními dosahuje hodnoty 54,1%, tj. první kritérium je splněno.*
2. *druhá největší komponenta, tj. první kontrast v reziduích, nebude větší než 5 % ze zbylého rozptylu nebo hodnota eigenvalue, tzv. vlastního čísla, bude < 2 – nevysvětlitelný rozptyl v 1. kontrastu dosahuje hodnoty 8,8%, tj. hodnota je vyšší než požadovaných 5%, ale hodnota eigenvalue v 1. kontrastu je 1,7225, tj. hodnota je menší než 2, to znamená, že jedna z požadovaných podmínek je splněna a tím je splněno i druhé kritérium.*
3. *minimální poměr 3:1 bude mezi rozptylem vysvětleným položkami a rozptylem vysvětleným prvním kontrastem, druhou největší komponentou – poměr dosahuje hodnoty 3,171:1 (27,9% : 8,8% = 3,171), tj. kritérium je splněno.*

5.5 Standardizace škály – reliabilita

Výsledky reliability jsou uvedeny v *tabulce č. 14*. Reliabilita byla počítána pomocí koeficientu KR – 20 a interpretována podle Tavakola a Dennicka (2011). Dále uvádíme v *tabulce č. 14* Raschovu reliabilitu pro testované osoby a položky a hodnoty separace u testovaných osob i položek. Ve výsledcích jsou také uvedeny hodnoty „reálné“ a „modelové“ reliability osob a položek.

Reliabilita posuzovací škály má hodnotu 0,81, což je podle Tavakola a Dennicka (2011) dobrý výsledek.

Reliabilita osob dosahuje nižších hodnot, hodnota 0,75 a hodnota separace testovaných osob 1,75 jsou nikoliv dobrým, ale podle Tavakola a Dennicka (2011) přijatelným výsledkem.

Reliabilita položek dosahuje hodnoty 0,99 a separační koeficient dosahuje hodnoty 8,77. Jedná se o velmi dobré výsledky a velmi vysokou reliabilitu.

Tabulka č. 14 Reliabilita škály

Reliabilita testu (Cronbach Alpha, KR-20)		0.81
	Separace testovaných osob	Reliabilita testovaných osob
Reálný RMSE	1.75	0.75
Modelový RMSE	1.91	0.78
	Separace položek	Reliabilita položek
Reálný RMSE	8.77	0.99
Modelový RMSE	9.11	0.99

5.6 Standardizace škály – objektivita

Objektivitu škály hodnotili čtyři examinátoři s různou délkou trenérských zkušeností (viz kapitola 6.4.5.). Objektivitu jsme počítali pro každou položku zvlášť a výsledné hodnoty včetně interpretace výsledků podle Landise a Kocha (1977) jsme zaznamenali do *tabulky č. 15*.

Výsledná shoda čtyř examinátorů na všech devíti položkách je 0,985313 % a průměrná hodnota Fleissova kappu u devíti výsledných položek má hodnotu 0,936887. Jedná se o velice dobrý výsledek a téměř dokonalou shodu.

Tabulka č. 15 Objektivita škály

POLOŽKA	Shoda v %	Fleissovo kappu	Interpretace (Landis & Koch, 1977)
8	0.982759	0.963046	téměř dokonalá shoda
10	1.00000	1.00000	téměř dokonalá shoda
9	1.00000	1.00000	téměř dokonalá shoda
16	1.00000	1.00000	téměř dokonalá shoda
5	0.942529	0.865741	téměř dokonalá shoda
4	0.977011	0.765657	dobrá shoda
15	1.00000	1.00000	téměř dokonalá shoda
3	0.965517	0.837535	téměř dokonalá shoda
7	1.00000	1.00000	téměř dokonalá shoda

6. DISKUSE

Výsledky Raschovy analýzy u sestavené Guttmanovy škály naznačují, že výsledná devítipoložková škála určená k testování florbalových dovedností u hráčů mladšího školního věku dosahuje úrovně validního, reliabilního a objektivního standardizovaného diagnostického nástroje.

Pro sestrojení škály bylo vybráno celkem třicet položek, z nichž bylo dvanáct položek expertní analýzou vyřazeno. Pro pilotní testování bylo vybráno celkem osmnáct položek s Lawshovým koeficientem 0,2 a vyšším. Po pilotním testování byly vyřazené další dvě položky „Driblink míčku přes širokou čáru – hokejový driblink“ a „Driblink míčku přes širokou čáru – florbalový driblink“ z důvodu obtížného dichotomického hodnocení „splnil – nesplnil“ v provedení probandů.

Výběr šestnácti položek, které 212 probandů testovalo během hlavního testování, se ukázal vhodným výběrem z hlediska obtížnosti položek. Položky pokryly kontinuum diagnostikovaného rysu v dostatečném rozpětí a s pravidelným škálovým rozložením takovým způsobem, že z šestnácti položek mohlo být vybráno devět položek pro výslednou škálu, která splňuje požadavky na diagnostický nástroj právě z hlediska pokrytí kontinua diagnostikovaného rysu v dostatečném rozpětí a s pravidelným škálovým rozložením.

Z výsledných devíti položek se nejlehčí položka č. 7 „Přehození míčku přes mantinel bekhendovou stranou čepele“ blížila hodnotě -3 logitů (-2.85 logitů) a naopak nejtěžší položka č. 8 „Vedení míčku po čáře“ na opačné straně kontinua škály mírně přesáhla hodnotu $+3$ logity ($3,01$ logitů). Pro účely práce jsme se rozhodli tuto hodnotu akceptovat. Položka č. 5 „Vyhození míčku z čepele a chycení zpět na čepel“ se uprostřed škály blížila hodnotě 0 ($0,01$ logitů). Ostatních šest položek pokrylo intervaly mezi hodnotami $(0,1)$ a $(0, -1)$, $(1,2)$ a $(-1, -2)$, $(2,3)$ a $(-2, -3)$ logitů. Požadavek na pravidelné rozložení škálových hodnot (Břicháček 1978, Čepička 2003, 2005) byl splněn.

Vyřazení sedmi položek z původního počtu šestnácti položek, které byly podstoupeny testování, bylo provedeno po Raschově analýze a následném vyhodnocení. Každou z šestnácti položek jsme zkoumali z několika hledisek – infit a outfit statistik a hodnot parametru obtížnosti. Parametr obtížnosti musí dosahovat hodnot od -3 do 3 logitů. Hodnoty ztrátových fit funkcí u středního kvadratického rezidua (MNSQ) by se měly pohybovat v intervalu

(0,5;1,5), u standardizované hodnoty rezidua (ZSTD) by se hodnoty měly v ideálním případě rovnat 0 (Linacre, 2017).

Po zhodnocení fit statistik jsme se rozhodli podle doporučení McCreary a kol. (2014) a Linacre (2017) nejprve pro vyřazení 4 probandů s nízkou úrovní diagnostikovaného rysu, kteří náhodně skórovali na obtížných položkách č. 8 a č. 10. Položky měly vyhovující hodnoty parametrů obtížnosti, ale příliš vysoké hodnoty outfit MNSQ. Po následující nové Raschově analýze výsledků testování zbylých 208 probandů jsme rozhodovali o zařazení nebo vyřazení položek do výsledné škály podle doporučení McCreary a kol. (2014) a Linacreho (2017). Položky č. 1, 2, 6, 11, 12, 13 a 14 byly vyřazeny a zbylých devět položek bylo zařazeno do výsledné škály, která byla standardizována.

Konstruktová (teoretická) validita škály byla posuzována na základě shody modelu s daty, která je u Raschova modelu posuzována fit statistikami (Čepička 2005). Hodnoty ztrátových fit funkcí u středního kvadratického rezidua (MNSQ) by se měly pohybovat v intervalu (0,5;1,5) (Linacre, 2017), stejně tak jako u posuzování předchozích 16-ti položek. Hodnoty infit a outfit MNSQ u všech devíti položek se pohybovaly v doporučeném intervalu, dokonce většina hodnot splňovala i přísnější interval (0,7;1,3) (Čepička 2003, 2005; Zhu 1996; Wilson 2005). Jediné tři hodnoty outfit MNSQ u položky č. 5 (0,53), položky č. 4 (0,66) a položky č. 7 (0,65) splňovaly mírnější podmínky podle Linacreho (2017).

Zkoumání shody modelu s daty bylo také prvním krokem k posuzování unidimenzionality škály, základního předpokladu Guttmanovy škály. Druhým krokem byla analýza hlavních komponent standardizovaných reziduí, při které jsme hodnotili tři kritéria, která byla splněna – první komponenta vysvětlovala 54,1% celkového rozptylu v datech, tj. více než požadovaných 50%, hodnota *eigenvalue* v prvním kontrastu reziduí byla 1,7225, tj. menší než 2, a poměr mezi rozptylem vysvětleným položkami a rozptylem vysvětleným prvním kontrastem byl 3,171:1, tj. více než 3:1. Jediná hodnota, která neodpovídala požadovaným kritériím, byla hodnota prvního kontrastu v reziduích, která neměla přesahovat hodnotu 5%, ale dosahovala hodnoty 8,8 %. Tento výsledek by mohl naznačovat existenci druhé dimenze (Linacre, 2017), ale protože výše uvedená hodnota *eigenvalue* v prvním kontrastu reziduí je nízká, měla by tím být existence druhé dimenze vyloučena a unidimenzionalita potvrzena.

Další důležitou vlastností standardizované škály je reliabilita. Reliabilita devítipoložkové škály dosáhla hodnoty 0,81, podle interpretace podle Tavakola a Dennicka (2011) se jedná o dobrou hodnotu.

Důležitá je i hodnota reliability položek, která dosahuje hodnoty 0,99 a separační koeficient pro reálné hodnoty dosahuje velikosti 8,77. Hodnoty naznačují, že testovaný soubor osob byl dostatečně velký pro potvrzení posloupnosti položek a jednalo se o dostatečný počet různě obtížných položek (Linacre, 2017).

Nižší hodnoty byly zjištěny u reliability testovaných osob – 0,75 a separace testovaných osob – 1,75. Hodnota testovaných osob nedosahuje hodnot $> 0,8$ a separační koeficient osob nedosahuje hodnot > 2 , což jsou požadované hodnoty při dobré reliabilitě. Tyto výsledky podle Linacreho (2017) naznačují, že u testovaných osob se neobjevil dostatečně velký rozsah testovaného rysu. Pro zvýšení reliability testovaných osob Linacre (2017) doporučuje testovat osoby s extrémními hodnotami měřeného rysu (s extrémně vysokou nebo nízkou úrovní dovednosti). Dále uvádí, že nižší reliabilitu může způsobovat nedostatečná délka testu nebo nesprávné skórování osob na odpovídajících položkách. Domníváme se, že hlavním důvodem nižší reliability testovaných osob je nedostatečný počet osob s extrémně vysokou nebo nízkou úrovní testované dovednosti. Když zhodnotíme úroveň testované dovednosti mezi testovanými osobami, zastoupení extrémních hodnot zde pravděpodobně chybí. Nicméně uvedené hodnoty můžeme podle Tavakola a Dennicka (2011) stále považovat za přijatelné.

Hodnota Fleissova koeficientu kappa, pomocí kterého jsme hodnotili objektivitu škály, dosáhla u osmi položek z devíti výsledných položek téměř dokonalé shody (hodnoty $> 0,81$). Jen u položky č. 4, zvednutí míčku na čepel, se jednalo o menší shodu ($\kappa = 0,765657$), která je ale podle interpretace Landise a Kocha (1977) dobrá. Dále měla nižší hodnotu kappa koeficientu položka č. 3, osmička ($\kappa = 0,837535$), položka č. 5, vyhození míčku z čepele a chycení zpět ($\kappa = 0,865741$) a položka č. 8, vedení míčku po čáře ($\kappa = 0,963046$). Jedná se o položky, které se hůře hodnotí, protože se nejedná o hodnocení typu zasažení cíle, měření času u slalomu nebo přehození překážky, které lze jednodušeji dichotomicky ohodnotit. Shoda examinátorů u položek tohoto typu dosáhla dokonalé shody $\kappa = 1$.

7. ZÁVĚR

V naší disertační práci jsme se pokusili prokázat, že přístupy kinantropologie mohou výrazně obohatit sportovní praxi, a to cestou aplikace vědeckých postupů standardizace diagnostických nástrojů. Pro přípravu dětí ve florbalu zatím neexistují žádné ověřené diagnostické nástroje (testy, posuzovací škály) pro posouzení florbalových dovedností v tréninku a často se využívají diagnostické nástroje bez příslušných vědeckých parametrů.

Výsledná škála obsahuje devět položek, které hodnotí úroveň manipulace s florbalovou holí a míčkem:

- přehození míčku přes mantinel bekhendovou stranou čepele,
- osmička,
- slalom s míčkem č. 2,
- zvednutí míčku na čepel,
- vyhození míčku z čepele a chycení zpět,
- slalom s míčkem č. 3,
- lob míčku do vymezeného prostoru,
- přihrávka z místa na cíl,
- vedení míčku po čáře.

Pozitivním výsledkem u posuzovací škály je pravidelné rozložení položek po celém kontinuu diagnostikovaného rysu, postupné zvyšování obtížnosti položek od nejjednodušší po nejobtížnější a relativně jednoduchý obsah položek umožňující praktickou aplikaci. Testování hráčů florbalu škálou o devíti položkách je jednoduše realizovatelné v tréninkovém procesu z hlediska časové náročnosti.

Vytvořená škála byla standardizována a výsledky ukázaly, že se jedná o validní, reliabilní a objektivní diagnostický nástroj. Nižší výsledné hodnoty se objevily pouze v případě Raschovy reliability testovaných osob. Interpretace těchto výsledků nám naznačila, že v testovaném souboru nebyly osoby s dostatečně velkým rozsahem testovaného rysu,

pravděpodobně chyběl dostatečný počet testovaných osob s extrémní, vysokou nebo nízkou, úrovní testované dovednosti.

Ve výsledné posuzovací škále chybí v praxi akcentovaná položka obsahující florbalovou střelbu (položky č. 12 a č. 13). Střelecké položky byly pro testovaný soubor příliš obtížné. I když jsme se snažili o rovnoměrné zastoupení všech úrovní diagnostikovaného rysu u probandů, při praktické realizaci bylo přítomno více probandů se střední nebo nižší úrovní florbalových dovedností, pro které byly střelecké položky příliš obtížné. Z tohoto hlediska by bylo vhodné 16 původních položek otestovat na hráčích s vyšší úrovní diagnostikovaného rysu za účelem posunutí obtížnosti škály směrem k obtížnějším položkám. Jinou možností by bylo zjednodušení střeleckých položek (například pro úspěšné splnění položky by stačil jen jeden úspěšný střelecký pokus ze tří celkových pokusů, zkrácení vzdálenosti od florbalové branky, apod.) pokud by byla možnost nového testování.

I přes absenci střeleckých položek však považujeme posuzovací škálu za velmi kvalitní diagnostický nástroj, který hodnotí manipulaci s míčkem pomocí florbalové hole. Diagnostický nástroj lze použít při florbalovém tréninku, na florbalových soustředěních a táborech, na kempech talentované mládeže a také jako diagnostický nástroj při výběru talentů. Domníváme se, že se jedná o velice vhodný nástroj určený pro florbalové trenéry, pomocí kterého lze sledovat úroveň dovedností u hráčů florbalu mladšího školního věku. Dále se domníváme, že pomocí opakovaného testování posuzovací škálou lze také sledovat postupný vývoj úrovně florbalových dovedností u hráčů florbalu. Tvrzení by mělo být potvrzené obsáhlejší vědeckým výzkumem, který by se mohl stát následujícím krokem v dalším vývoji tohoto diagnostického nástroje.

Z pohledu kinantropologie jako vědeckého oboru lze považovat za přínosný popsání přehledný postup při konstrukci posuzovací škály, dále jsou popsány jednotlivé kroky při standardizaci vědeckého nástroje a také lze v práci najít přehled možností při interpretaci výsledků. Z praktického hlediska má florbal pro zkoumanou věkovou kategorii ověřenou standardizovanou posuzovací škálu, avšak postup při konstrukci tohoto diagnostického nástroje lze použít i u ostatních sportovních specializací s cílem vytvořit vhodný standardizovaný nástroj určený k testování úrovně sportovních dovedností u daného sportu.

Výsledky práce jsou určeny pro trenéry mládežnických kategorií ve florbalu. Výzkumná práce by měla být také základem pro další metodický materiál České florbalové unie, který by zvýšil dosah a využitelnost těchto informací pro co nejširší okruh trenérů florbalu v České republice.

8. LITERATURA

1. ADAMÍROVÁ, J. a kol. *Psychomotorika*. Praha: ČSAPV, 2010. ISBN 8086586243.
2. ANDRICH, D. *Rasch models for measurement*. Newbury park: Sage Publications, 1988. ISBN 0803936478.
3. ATKINSON, R. C. a kol. *Psychologie*. Praha: Portál, 2003. ISBN 8071786403.
4. BAKER, F. B., KIM S. H. *Item Response Theory. Parametr Estimation Techniques*. New York: Marcel Dekker, 2004. ISBN 0824758250.
5. BALYI, I. a kol. *Canadian sport for life: Long-term athlete development resource paper*. Vancouver: Canadian Sport Centres, 2005.
6. BALYI, I., ROSS, G. Key coaching concerning growth and maturation of the young developing performer. In: Balyi, I., Williams, C. *Coaching the young developing performer*. Leeds: Coachwise, 2009, s. 39-45.
7. BASÍLIO, M. L. Cross – cultural validity of the Brazilian version of the Abilhand questionnaire for chronic stroke individuals, based on Rasch analyses. *J Rehabil Med*, 2016, roč. 48, s. 6-13.
8. BAUMGARTNER, T. A., JACKSON, A. S., MAHAR, M. T., ROWE, A. D. *Measurement for evaluation in physical education and exercise science*. Boston: McGraw Hill, 2003. ISBN 0-07-246183-7.
9. BELEJ, M. *Motorické učenie*. Prešov: Prešovská univerzita, 2001. ISBN 8080680418.
10. BERKA, K. *Měření: pojmy, teorie, problémy*. Praha: Academia, 1977.
11. BLAHUŠ, P. Matematická teorie škálování a možnosti jejího využití v tělesné výchově – přehled nových poznatků. *Acta Universitatis Carolinae Gymnica*, 1981, roč. 17, č. 2, s. 81-96.
12. BLAHUŠ, P. Faktorová analýza a její zobecnění. Praha: SNTL, 1985.
13. BLAHUŠ, P. *K systémovému pojetí statistických metod v metodologii empirického výzkumu chování*. Praha: Karolinum, 1996. ISBN 80-7184-100-5.
14. BOUCHARD, C., WIECZOREK, E. *Problems of sports medicine and of sports training and coaching*. S. I.: 1975.
15. BŘICHÁČEK, V. *Úvod do psychologického škálování*. Bratislava: Psychodiagnostické a didaktické testy, 1978.
16. BUKAČ, L., DOVALIL, J. *Inovace ledního hokeje: technicko-taktická příprava*. Praha: Ústřední výbor ČSTV, 1989.

17. BUKAČ, L., DOVALIL, J. *Lední hokej: trénink herní dokonalosti*. Praha: Olympia, 1990. ISBN 80-7033-024-4.
18. BUKAČ, L. *Intelekt, učení, dovednosti a koučování v ledním hokeji: komprehenzivní pohled na utkání, trénink a rozvoj individuálního herního výkonu*. Praha: Olympia, 2005. ISBN 80-7033-896-2.
19. BUKAČ, L. *Koučování mládeže ledního hokeje*. Praha: Český svaz ledního hokeje, 2012.
20. BURTON, A. W., MILLER, D. E. *Movement skill assessment*. Champaign: Human Kinetics, 1998. ISBN 9780873229753.
21. CLEEREMANS, A. Models of implicit learning. In: NADAL, L. (ed.). *Encyclopedia of cognitive sciences*. London: Macmillan Publishers, 2002, s. 491-499.
22. CORTINA, J. M. What Is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, 1993, roč. 78, č. 1, s. 98-104.
23. CRONBACH, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika*, 1951, roč. 16, č. 3, s. 297-334.
24. ČELIKOVSKÝ, S. a kol. *Antropomotorika: pro studující tělesnou výchovu*. Praha: SPN, 1979.
25. ČEPIČKA, L. Stanovení obtížnosti motorického testu. *Česká kinantropologie*, 1999, roč. 3, č. 1, s. 87-94.
26. ČEPIČKA, L. *Modely teorie položkových odpovědí v diagnostice motoriky člověka*. Plzeň: Západočeská univerzita, 2002.
27. ČEPIČKA, L. Konstrukce perfektní škály v diagnostice motorických dovedností. *Česká kinantropologie*, 2003, roč. 7, č. 1, s. 7-18.
28. ČEPIČKA, L. The Rasch model in the motor ability testing. *Journal of Human Movement Studies*, 2003, roč. 10, s. 99-106.
29. ČEPIČKA, L. Stepnicka's modification of the Brace test: An investigation of dimensionality. *Perceptual and Motor Skills*, 2004, roč. 98, s. 171-178.
30. ČEPIČKA, L. *Příspěvek k unidimenzionálnímu škálování motorických předpokladů*. Praha, 2005. 181 s. Habilitační práce na FTVS UK.
31. ČEPIČKA, L. Normative data for the test of gross motor development-2 in 7-yr.-old children in the Czech Republic. *Perceptual and Motor Skills*, 2010, roč. 110, s. 1048-1052.
32. DOBRÝ, L., VELENSKÝ, E. *Košíková teorie a didaktika*. Praha: SPN, 1987.
33. DONOVAN, N. J. Adding meaning to measurement: Initial Rasch analysis of the ASHA FACS Social Communication Subtest. *Aphasiology*, 2006, roč. 20, s. 362-373.

34. DOVALIL, J. a kol. *Lexikon sportovního tréninku*. Praha: Karolinum, 2008. ISBN 978-80-246-1404-5.
35. DOVALIL, J. a kol. *Výkon a trénink ve sportu*. Praha: Olympia, 2012. ISBN 80-7033-760-5.
36. DRAGOUNOVÁ, Z., PERIČ, T., DOVALIL, J. Implicitní motorické učení – možnosti ve sportovním tréninku. *Česká kinantropologie*, 2013, roč. 17, č. 3, s. 11-22.
37. FETZ, F. *Leibesübungen für 6- bis 15 jährige: praktischer Lehrbehelf für Leibeserzieher*. Wien: Österreichischer Bundesverlag, 1975. ISBN 3-215-50504-5.
38. FLEISS, J. L. Measuring nominal scale agreement among many raters. *Psychol Bull*, 1971, roč. 76, s. 378–82.
39. GAVORA, P. *Úvod do pedagogického výzkumu*. Brno: Paido, 2010. ISBN 8073151855.
40. GRAFFIGNA, G. a kol. Measuring patient activation in Italy: Translation, adaptation and validation of the Italian version of the patient activation measure 13 (PAM13-I). *BMC Medical Informatics and Decision Making*, 2015, roč. 15, s. 109.
41. HENDL, J. *Přehled statistických metod: analýza a metaanalýza dat*. Praha: Portál, 2015. ISBN 978-80-7367-482-3.
42. HENDL, J. *Kvalitativní výzkum: základní teorie, metody a aplikace*. Praha: Portál, 2016. ISBN 978-80-2620-982-6.
43. CHOUTKA, M. *Studium struktury sportovních výkonů*. Praha: Univerzita Karlova, 1976.
44. CHOUTKOVÁ, B., KUČERA, M. *Mládež a sport*. Praha: Olympia, 1970.
45. CHRÁSKA, M. *Metody pedagogického výzkumu: základy kvantitativního výzkumu*. Praha: Grada Publishing, 2007. ISBN 802475326X.
46. JANSÁ, P. *Využití škálovacích technik při posuzování pohybových dovedností v tělesné výchově*. Praha, 1975. Kandidátská disertační práce na FTVS UK.
47. JANSÁ, P. Mnohorozměrové škálování v tělesné výchově a sportu. *Teorie a praxe tělesné výchovy*, 1989, roč. 37, č. 2.
48. JANSÁ, P. a kol. *Pedagogika sportu*. Praha: Karolinum, 2012. ISBN 802462026X.
49. KARLBERG, M. *Teknikmärke*. [online]. Solna : Svenska Innebandyförbundet, 2013. [cit. 2017-10-10]. Dostupné z: http://www.innebandy.se/Global/SIBF/Forbundsinfo/Utveckling/Teknikm%C3%A4rket/Teknikm%C3%A4rke%20bl%C3%A5_korr.pdf.
50. KAVALÍŘOVÁ, G. Aplikace škál při odborném posuzování pohybových dovedností. *Česká kinantropologie*, 2003, roč. 7, č. 2, s. 79-90.

51. KIRCHNER, G. Physical education for elementary school children. W. C.: Brown, 1981. ISBN 0697071758.
52. KLENEROVÁ, V., HYNIE, S. Paměť a její poruchy. *Čs. fyziologie*, 2010, roč. 59, č. 1, s. 15-20.
53. KNUDSON, D. V., MORRISON, C. S. *Qualitative analysis of human movement*. Champaign: Human Kinetics, 2002. ISBN 0-7360-3462-5.
54. KODÝM, M. *Model úspěšného hráče*. In: *Odbíjená – Košíková*, 1969, roč. 19, č. 2, s. 17.
55. KOSTKA, V. *Lední hokej*. Praha: Státní tělovýchovné nakladatelství, 1955.
56. KOSTKA, V., BUKAČ, L., ŠAFAŘÍK, V. *Lední hokej: teorie a didaktika*. [online]. Praha: Státní pedagogické nakladatelství, 1986 [cit. 2017-10-27]. Dostupné z: <http://alephuk.cuni.cz/CKIS-28.html>.
57. KOUKOLÍK, F. *Mozek a jeho duše*. Praha: Galén, 2008. ISBN: 8074920690.
58. KYSEL, J. *Florbal – kompletní průvodce*. Praha: ČFbU, 2010. ISBN: 978-80-247-3615-0.
59. LAWSHE, C. H. A quantitative approach to content validity. *Personnel psychology*, 1975, roč. 28, s. 563-575.
60. LANDIS, J. R., KOCH, G. G. The measurement of observer agreement for categorical data. *Biometrics*, 1977, roč. 33, č. 1, s. 159–74.
61. LIAO, C., MASTERS, R. S. W. Analogy learning: A means to implicit motor learning. *Journal of Sports Sciences*, 2001, roč. 19, s. 307-319.
62. LINACRE, J. M. *Winsteps® Rasch measurement computer program manual*. [online]. Beaverton: Winsteps, 2017. [cit. 2017-9-5]. Dostupné z: <http://www.winsteps.com>.
63. MARTÍNKOVÁ, Z. *Florbal – praktický průvodce tréninkem mládeže*. Praha: ČFbU, 2009.
64. MASTERS, R. S. W. Knowledge, knerves and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology*, 1992, roč. 83, s. 43-359.
65. MASTERS, R. S. W., MAXWELL, J. P. Implicit motor learning, reinvestment and movement disruption: What you don't know won't hurt you? In: WILLIAMS, A. M., HODGES, N. J. *Skill acquisition in sport: Research, theory and practice*. London: Routledge, 2004, s. 207-228.
66. MASSOF, W. Understanding rasch and Item response theory Models: Applications to the Estimation and Validation of Interval Latent trait Measures from responses to rating Scale Questionnaires. *Ophthalmic Epidemiology*, 2011, roč. 18, č. 1, s. 1-19.

67. MAXWELL, J. P., MASTERS, R. S. W., EVES, F. F. From novice to no know-how: A longitudinal study of implicit motor learning. *Journal of Sports Science*, 2000, roč. 18, s. 111-120.
68. MAXWELL, J. P. Skill learning the implicit way. In: FARROW, D., BAKER, J., MACMAHON, C. *Developing sport expertise*. Abingdon: Routledge, 2008, s. 92-101.
69. MECNER, J. Příspěvek k hodnocení spolehlivosti testu házení na cíl. *Tělesná výchova a sport mládeže*, 1975, roč. 23, s. 727-733.
70. MĚKOTA, K. *Měření a testy v antropomotorice (I. a II. díl)*. Olomouc: Univerzita Palackého, 1973.
71. MĚKOTA, K., BLAHUŠ, P. *Motorické testy v tělesné výchově*. Praha: SPN, 1983. ISBN 14-467-83.
72. MĚKOTA, K., CUBEREK, R. *Pohybové dovednosti, činnosti, výkony*. Olomouc: Univerzita Palackého, 2007. ISBN 978-80-244-1728-8.
73. MIKAČOVÁ, M. *Sledování herního výkonu ve florbalu*. Brno, 2008. Rigorózní práce na FSS MU.
74. MORROW, J. R., JACKSON, A. W., DISCH, J. G., MOOD, D. P. *Measurement and evaluation in human performance*. Champaign: Human Kinetics, 2005. ISBN 0-7360-5540-1.
75. MYERS, N. D., FELTZ, D. L., WOLFE, W. A Confirmatory Study of Rating Scale Category Effectiveness for the Coaching Efficacy Scale. *Research Quarterly for Exercise and Sport*, 2008, roč. 79, č. 3, s. 300-311.
76. NEUMAN, J. *Cvičení a testy obratnosti, vytrvalosti a síly*. Praha: Portál, 2003. ISBN 8071787302.
77. OSTINI, R., NERING, M. L. *Polytomous Item Response Theory Models*. Thousand Oaks: Sage Publications, 2006. ISBN 076193068X.
78. PAAVILAINEN, A. *Individual Technique and Tactics*. IFF, 2007.
79. PAVLÍČEK, M. *Guttmanova škála pro výběr sportovních talentů ve volejbalu*. Plzeň, 2003. *Diplomová práce na PF ZU*.
80. PAVLIŠ, Z. a kol. *Školení trenérů ledního hokeje*. Praha: ČSLH, 2000.
81. PAVLIŠ, Z., PERIČ, T., NOVÁK, Z., BERÁNEK, J. *Příručka pro trenéry ledního hokeje I.část*. Praha: ČSLH, 1998.
82. PAVLIŠ, Z., PERIČ, T., NOVÁK, Z., MAZANEC, M. *Příručka pro trenéry ledního hokeje II.část*. Praha: ČSLH, 2000.
83. PAVLIŠ, Z. a kol. *Příručka pro trenéry ledního hokeje III.část*. Praha: ČSLH, 2002.

84. PERIČ, T. *Lední hokej*. Praha: Grada, 2002. ISBN 8024704722.
85. PERIČ, T. *Sportovní příprava dětí*. Praha: Grada, 2012. ISBN 978-80-247-2643-4.
86. POOLTON, J. M., ZACHRY, T. L. So you want to learn implicitly? Coaching and learning through implicit motor learning techniques. *International Journal of Sports Science & Coaching*, 2007, roč. 2, č. 1, s. 67-78.
87. RANDOLPH, J. J. *Online Kappa Calculator computer software*. [online]. [cit. 2018-01-18]. Dostupné z: <http://justus.randolph.name/kappa>.
88. RYCHTECKÝ, A., FIALOVÁ, L. *Didaktika školní tělesné výchovy*. Praha: Karolinum, 2002. ISBN: 80-7184-127-7.
89. SABARI J. S. a kol. Rasch Analysis of a New Hierarchical Scoring System for Evaluating Hand Function on the Motor Assessment Scale for Stroke. *Stroke Research and Treatment Volume*, 2014, Article ID 730298.
90. SAFRIT, M. J. The Applicability of Item Response Theory to Test of Motor Behavior. *Research Quarterly for Exercise and Sport*, 1987, roč. 58, s. 213-215.
91. SAFRIT, M. J. Item response theory and measurement of motor behavior. *Research Quarterly for Exercise and Sport*, 1989, roč. 60, s. 325-335.
92. SAFRIT, M. J., ZHU, W., COSTA, M. G., ZHANG, L. The difficulty of sit-ups test: An empirical investigation. *Research Quarterly for Exercise and Sport*, 1992, roč. 63, s. 277-283.
93. SCHMIDT, R. A. *Motor learning and performance: a problem-based learning approach*. Champaign: Human Kinetics, 2004. ISBN 0-7360-4566-X.
94. SCHMIDT, R. A. *Motor learning and performance: from principles to practice*. Champaign: Human Kinetics, 1991. ISBN 0-87322-308-X.
95. SCHMIDT, R. A., WRISBERG, C. A. *Motor learning and performance: A Situation-Based Learning Approach*. Champaign: Human Kinetics, 2008. ISBN : 0880115009.
96. SKRONDAL, A., RABE-HESKETH, S. *Generalized Latent Variable Modeling*. Boca Raton: Chapman&Hall/CRC, 2004. ISBN 1584880007.
97. SKRUŽNÝ, Z. a kol. *Florbal*. Praha: Grada, 2005. ISBN: 8024703831.
98. SMITH, E. V. a kol. *Using Item Mean Squares to Evaluate Fit to the Rasch Model*. [online]. A paper presented at the 1995 Annual Meeting of the American Educational Research Association in san francisco. [cit. 2018-1-25]. Dostupné z: <https://files.eric.ed.gov/fulltext/ED384617.pdf>.
99. SMITH, E. V. Using item mean squares to evaluate fit to the Rasch model. *J Outcome Meas.*, 1998, roč. 2, č. 1, s. 66-78.

100. SMITH, E. V. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, 2002, roč. 3, s. 205-231.
101. DE SOUZA L. H. The Development of an Unidimensional Scale of the Guttman Type for the Assessment of Mobility Disability in Multiple Sclerosis. *Clin Rehabil*, 1999, roč. 13, s. 476-481.
102. SPRAY, J. Recent developments in measurements and possible applications to the measurement to psychomotor behavior. *Research Quarterly for Exercise and Sport*, 1987, roč. 58, s. 203-209.
103. SPRAY, J. One-parametr item response theory models for psychomotor tests involving repeated, independent attempts. *Research Quarterly for Exercise and Sport*, 1990, roč. 61, s. 162-168.
104. ŠŮSS, V. *Význam indikátorů herního výkonu pro řízení tréninkového procesu*. Praha: Karolinum, 2006. ISBN 80-246-1162-7.
105. ŠIMON, J. *Struktura sportovního výkonu v hodě oštěpem: dynamika růstu tréninkového zatížení oštěpařů*. Praha: Ústřední výbor ČSTV, 1988.
106. ŠTILEC, M. *Sportovní příprava dětí a mládeže: pro posl. fak. tělesné výchovy a sportu*. Praha: SPN, 1989. ISBN 8070660260.
107. ŠTOCHL, J., MUSÁLEK, M. A practical guide to pilot standardization of tests. *Acta Universitatis Carolinae Kinanthropologicae*, 2009, roč. 45, č. 2, s. 5-15.
108. ŠVEC, Š. a kol. *Metodologie věd o výchově*. Brno: Paido, 2009. ISBN 8073151928.
109. TAVAKOL, M., DENNICK, R. Making Sense of Cronbach's Alpha. *International Journal of Medical Education*, 2011, roč. 2, s. 53-55.
110. TENENBAUM, G. The implementation of Thurstone's and Guttman's measurement ideas in Rasch analysis. *Int. J. Sport Psychol.*, 1999, roč. 30, s. 3-16.
111. THOMAS, J. R., NELSON, J. K., SILVERMAN, S. J. *Research methods in physical activity*. Champaign: Human Kinetics, 2005. ISBN 0-7360-5620-3.
112. TURNER-STOKES, L. The Work-ability Support Scale: Evaluation of Scoring Accuracy and Rater Reliability. *J Occup Rehabil*, 2014, roč. 24, s. 511-524.
113. TŮMA, A. *Testování motorických dovedností ve florbale*. Liberec, 2004. Diplomová práce na PF TU.
114. URBÁNEK, T., ŠIMEČEK, M. Teorie odpovědi na položku. *Československá psychologie*, 2001, roč. 45, s. 428-440.

115. VÁGNER, M. Využití vícerozměrného škálování a motorických testů pro hodnocení a výběr vojáků do kurzů boje zblízka. Praha, 2010. 158 s. Disertační práce na FTVS UK.
116. VELENSKÝ, M. *Pojetí basketbalového učiva pro děti a mládež*. Praha: Karolinum 2008. ISBN 978-80-246-1480-9.
117. VELOZO, C. A. a kol. Utilizing Rasch measurement models to develop a computer adaptive self-report of walking, climbing, and running. *Disability and Rehabilitation*, 2008, roč. 30, č. 6, s. 458-467.
118. WEIS, W., WEIS, U.: Beitrag zur Theorie der Leistung im Sport. In: CHOUTKA, M.: *Studium struktury sportovních výkonů*. Praha: Univerzita Karlova, 1976.
119. WELCH, V. Open Access Systematic reviews need to consider applicability to disadvantaged populations: inter-rater agreement for a health equity plausibility algorithm. *Medical Research Methodology*, 2012, roč. 12, s. 187.
120. VAN DE WINCKEL, A. a kol. Can quality of movement be measured? Rasch analysis and inter-rater reliability of the Motor Evaluation Scale for Upper Extremity in Stroke Patients (MESUPES). *Clinical Rehabilitation*, 2006, roč. 20, č. 10, s. 871-884.
121. WILSON, M. *Constructing measures: An item response modeling approach*. Mahwah, NJ, Erlbaum Associates, 2005.
122. WOOD, T. Putting Item Response Theory Into Perspective. *Research Quarterly for Exercise and Sport*, 1987, roč. 58, s. 216-220.
123. WRIGHT, B. D. Local dependency, correlations and principal components. *Rasch Meas Tran*, 1996, roč. 10, s. 509-511.
124. WRIGHT, B. D., MASTERS, G. N. *Rating scale analysis*. Chicago Illinois: Mesa Press, 1982.
125. WU, T. Y. Rasch Analysis of the General Self-Efficacy Scale in Workers with Traumatic Limb Injuries. *J Occup Rehabil*, 2016, roč. 26, s. 332-339.
126. ZAPF, A. a kol. Measuring inter-rater reliability for nominal data – which coefficients and confidence intervals are appropriate? *Medical Research Methodology*, 2016, roč. 16.
127. ZHÁNĚL, J. *Aplikace výzkumných metod v kinantropologii*. Brno: Masarykova univerzita, 2014. ISBN 978-80-210-6792-9.
128. ZHU, W., COLE, E. L. Many-Faceted Rasch Calibration of a Gross Motor Instrument. *Research Quarterly for Exercise and Sport*, 1996, roč. 67, č. 1, s. 24-34.
129. ZHU, W. Should total scores from a rating scale be used directly? *Research Quarterly for Exercise and Sport*, 1996, roč. 67, s. 363-372.

130. ZHU, W., KANG, S. J. Cross – Cultural Stability of the Optimal Categorization of a Self-Efficacy Scale: A Rasch Analysis. *Measurement in Physical Education and Exercise Science*, 1998, roč. 2, č. 4, s. 225-241.
131. ZHU, W. Constructing Tests Using Item Response Theory. In: WOOD, T. M., ZHU, W. *Measurement Theory And Practice in Kinesiology*. Champaign: Human Kinetics, 2006. ISBN 9780736045032.
132. ZHU, W. a kol. PE Metrics: Background, Testing Theory and Methods. *Research Quarterly for Exercise and Sport*, 2011, roč. 15, č. 2, s. 87-99.
133. ZHU, W. a kol. Development and Calibration of an Item Bank for PE Metrics Assessments: Standard 1. *Research Quarterly for Exercise and Sport*, 2011, roč. 15, č. 2, s. 119-137.
134. ZLATNÍK, D. *Florbalový trénink v praxi, herní činnosti jednotlivce*. Praha: ČFbU, 2004.