

UNIVERZITA KARLOVA V PRAZE

Fakulta tělesné výchovy a sportu

DISERTAČNÍ PRÁCE

Mgr. Zuzana Dragounová

2018

UNIVERZITA KARLOVA V PRAZE

Fakulta tělesné výchovy a sportu

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

Vedoucí disertační práce:

Doc. PhDr. Josef Dovalil, CSc.

Vypracovala:

Mgr. Zuzana Dragounová

Praha 2018

Prohlašuji, že jsem disertační práci vypracovala samostatně pod vedením Doc. PhDr. Josefa Dovalila, CSc., a že jsem uvedla všechny použité literární a odborné zdroje.

V Praze dne

.....

Mgr. Zuzana Dragounová

Poděkování

Autorka vyjadřuje poděkování Doc. PhDr. Josefu Dovalilovi, CSc. za odborné vedení práce a možnost využití jeho znalostí a zkušeností. Dále bych chtěla poděkovat všem, kteří mě podporovali ve studiu, především celé své rodině.

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.

Evidenční list

Souhlasím se zapůjčením své disertační práce ke studijním účelům. Uživatel svým podpisem stvrzuje, že tuto disertační práci použil ke studiu a prohlašuje, že ji uvede mezi použitými prameny.

Jméno a příjmení:

Fakulta / katedra:

Datum vypůjčení:

Podpis:

Obsah

1.	ÚVOD	11
2.	TEORETICKÝ A METODOLOGICKÝ ZÁKLAD ŘEŠENÍ	13
2.1	Struktura florbalového výkonu	13
2.2	Pohybová dovednost	16
2.2.1	<i>Definice a charakteristika pohybové a sportovní dovednosti</i>	16
2.2.2	<i>Motorické učení</i>	18
2.3	Mladší školní věk a jeho charakteristika	22
2.4	Diagnostika pohybových dovedností.....	24
2.4.1	<i>Kvalitativní a kvantitativní výzkum</i>	25
2.4.2	<i>Škálování</i>	27
2.4.3	<i>Definice a vlastnosti motorického testu</i>	29
2.4.4	<i>Testy a testové baterie</i>	31
3.	GUTTMANOVA ŠKÁLA.....	37
4.	TEORIE POLOŽKOVÝCH ODPOVĚDÍ.....	41
4.1	Využití teorie položkových odpovědí.....	43
4.1.1	<i>Základní předpoklady použití teorie položkových odpovědí</i>	43
4.1.2	<i>Funkce položkových odpovědí</i>	44
4.2	Modely IRT	47
4.2.1	<i>Jednparametrické IRT modely a Raschův model</i>	48
4.2.2	<i>Dvoupametrické IRT modely</i>	52
4.2.3	<i>Tříparametrické IRT modely</i>	53
4.3	Aplikace modelů IRT v kinantropologii.....	54
5.	CÍL PRÁCE A ÚKOLY PRÁCE	56
6.	DESIGN A METODOLOGIE VÝZKUMU	58

6.1	Vymezení teoretické oblasti výzkumu	58
6.2	Výzkumný soubor.....	61
6.3	Výzkumné metody.....	63
6.3.1	<i>Expertní analýza pro posouzení obsahové validity podle Lawshe.....</i>	63
6.3.2	<i>Ztrátové fit funkce pro určení shody modelu s daty</i>	64
6.3.3	<i>Reliabilita škály u Raschova modelu</i>	65
6.3.4	<i>Fleissovo kappa pro zjištění shody posuzovatelů</i>	67
6.3.5	<i>Analýza hlavních komponent pro určení unidimenzionality škály.....</i>	69
6.4	Postup při konstrukci motorické škály	71
6.4.1	<i>Příprava výzkumu</i>	72
6.4.2	<i>Specifikace teoretického konceptu</i>	73
6.4.3	<i>Výběr položek.....</i>	73
6.4.4	<i>Ověření položek při pilotním testování</i>	76
6.4.5	<i>Hlavní sběr dat.....</i>	76
6.4.6	<i>Výpočet obtížnosti položek a sestavení škály</i>	77
6.4.7	<i>Posouzení validity, reliability, objektivity a unidimenzionality škály.....</i>	78
6.4.8	<i>Zhodnocení kvality škály.....</i>	83
7.	VÝSLEDKY	84
7.1	Výsledky expertní analýzy	84
7.2	Výsledky pilotního testování	88
7.3	Sestavení škály	90
7.3.1	<i>1. fáze – vyřazení nevyhovujících položek</i>	90
7.3.2	<i>2. fáze – kvantifikace obtížnosti jednotlivých položek</i>	93
7.3.3	<i>3. fáze – zhodnocení rozložení škálových hodnot</i>	94
7.4	Standardizace škály – validita	96
7.5	Standardizace škály – reliabilita	98
7.6	Standardizace škály – objektivita	99
8.	DISKUSE.....	100

9. ZÁVĚR	106
10. LITERATURA	108
11. PŘÍLOHY	117

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 přidavných 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 A METODOLOGICKÝ ZÁKLAD ŘEŠENÍ

2.1 Struktura florbalového výkonu

Pro potřeby účinného tréninku je nezbytné věnovat pozornost obsahu a průběhu pohybové činnosti. Analýza této činnosti má pro pochopení sportovního výkonu mimořádný význam.

Sportovní výkony se realizují ve specifických pohybových činnostech, jejichž obsahem je řešení úkolů, které vymezují pravidla příslušného sportu. Pohybové činnosti, které jsou ovlivněné vnějšími podmínkami, znamenají určité požadavky na jedince. Jedná se o komplexní integrovaný projev mnoha tělesných a psychických funkcí člověka.

Každý sportovní výkon – z hlediska jeho struktury – charakterizuje jak počet, tak i uspořádání faktorů. V některých výkonech může dominovat převážně jeden faktor (monofaktorální sportovní výkony), jiné jsou postaveny na existenci většího zastoupení faktorů (sportovní výkony multifaktorální) (Dovalil a kol., 2012).

Florbal jako ostatní kolektivní sporty se řadí mezi multifaktoriální sportovní výkony. Jednotlivé faktory vychází ze somatických, kondičních, technických, taktických a psychických základů výkonů.

Florbal patří mezi invazní – brankové sportovní hry, které hrají dvě družstva proti sobě na obdélníkovém hřišti s mantinely a zaoblenými rohy. Herním úkolem florbalového utkání je úderem florbalové hole dopravit míček do branky vícekrát než soupeř (Süss, 2006). Družstvo, které má v držení míček, plní útočné úkoly, tj. snaží se dopravit míček do branky soupeře. Soupeř, který nemá v držení míček, plní obranné úkoly, tj. brání vstřelení branky a snaží se získat míček. Florbalové utkání je limitované časem, který je rozdělen do tří dvacetiminutových úseků. Pro florbalové utkání jsou charakteristické neustále se měnící podmínky herní situace. Proměnlivé situace, ke kterým na hřišti dochází, bychom mohli rozdělit na úseky začínající a končící signálem rozhodčích. Rozhodčí zahajují hru signálem a opět ji signálem přerušují, pokud dojde ke vstřelení branky, porušení pravidel, vyhození míčku mimo hřiště, ukončení časového limitu apod. Takto dlouhé úseky můžeme ještě rozdělit na kratší proměnlivé situace jako například situace od rozehrání míčku po jeho ztrátu

z pohledu útočícího týmu. Následuje proměnlivá situace od ztráty míčku k jeho opětovnému získání (z pohledu útočícího týmu). Tyto proměnlivé situace se neustále opakují až do přerušení hry rozhodčími. Ve florbalovém utkání může dojít ještě k speciálním situacím v případě nerozhodného utkání. Pokud florbalový zápas není rozhodnut v základní hrací době ani v nastaveném čase, následují samostatné nájezdy. Samostatné nájezdy zahajuje hráč pohybem míčku na středu hřiště a situace končí vystřelením míčku na branku.

Proč je důležité zkoumat strukturu sportovního výkonu jednotlivého hráče v kolektivním sportu, tj. jakým způsobem se odráží sportovní výkon hráče ve hře celého týmu, je vyjádřeno u studia výkonu basketbalového hráče: „Sportovní výkon basketbalisty chápeme jako vývojový stupeň způsobilosti participovat (podílet se) na hře družstva. Je podmíněn komplexní dispoziční strukturou, která se navenek projevuje jako souhrn osvojených herních činností integrovaných do hry celého družstva“ (Dobry a Velenský, 1987).

Komplexním studiem sportovních výkonů se začala sportovní věda zabývat v 60. letech minulého století. Mezi autory, u kterých se poprvé objevovalo komplexnější chápání sportovního výkonu patří u nás například Kodým (1969), v Německu Fetz (1971), ve Švýcarsku W. Weiss a U. Weiss (1971) nebo v Kanadě kolektiv odborníků kolem Boucharda (1975).

Později se u nás studiem a výzkumem struktury sportovního výkonu věnoval především Choutka (1976), který se mimo individuálních sportů zabýval i strukturou sportovního výkonu fotbalistů. Výkonu hráče ledního hokeje se v Čechách věnovali na konci minulého století Kostka (1955, 1986), Bukač (2005), Bukač a Dovalil (1989, 1990), později Pavliš (1998, 2000, 2002) nebo Perič (2002). V souvislosti s basketbalovým výkonem se uvádí Dobry a Velenský (1987), v současné době M. Velenský (2008). V individuálních sportech vyšla v roce 1988 zevrubná publikace J. Šimona „Struktura sportovního výkonu v hodů oštěpem“.

Ve florbalu studoval herní výkon hráče Zlatník (2004), Skružný (2005), Kysel (2010), ze zahraničních autorů například Paavilainen (2007), která připravila metodický materiál pro Mezinárodní florbalovou unii. Ve své rigorózní práci se sledováním herního výkonu ve florbalu zabývala Mikačová (2008).

Podle vlastních zkušeností z dřívějších publikací (Martínková, 2009) by hráč florbalu měl být vyšší postavy s vyšším poměrem délky paží vzhledem k výšce a s odpovídající hmotností, která by neomezovala rychlý pohyb na hřišti. Zároveň hráč florbalu potřebuje

dostatek aktivní svalové hmoty pro stabilitu v osobních soubojích, vyšší úroveň silových schopností dolních končetin pro rychlý běh a změny směru a horních končetin pro manipulaci s florbalovou holí a střelbu. Somatotyp hráče by měl odpovídat mezomorfovi s vyrovnanou endoektomorfní složkou s mírnou převahou rychlých svalových vláken. Hodnota VO_2 max. by se měla pohybovat okolo 60 ml/kg.min. Hráč by měl mít vysokou úroveň rychlostních a koordinačních schopností a co nejširší rejstřík osvojených florbalových dovedností.

2.2 Pohybová dovednost

2.2.1 Definice a charakteristika pohybové a sportovní dovednosti

„Pohybové dovednosti jsou učením získané předpoklady sportovce správně, účelně, efektivně a úsporně řešit pohybové úkoly“ (Dovalil a kol., 2008). Podle Guthrie (Guthrie 1952 v Schmidt a Wrisberg, 2008, 9) spočívá mistrovství v předváděných dovednostech „ve schopnosti vytvoření konečných výstupů s maximální jistotou a minimálním úbytkem energie, nebo času a energie.“ Na základě Guthrieho charakteristiky vytvořil Schmidt (Schmidt a Wrisberg, 2008) tři základní kvalitativní rysy pro dosažení co nejvyšší úrovně předváděných dovedností:

- 1) *Maximální přesnost v dosažení cíle* – znamená dosahování vytyčeného cíle nebo výsledku s maximální přesností, opakovaně a bez vlivu náhody. Nejlepší sportovci dosahují výsledků na co nejvyšší úrovni dlouhodobě navzdory měnícím se okolnostem.
- 2) *Minimální ztráta energie* – jedná se o eliminaci nechtěných pohybů, které jsou neekonomické a nevedou k dosažení cíle, nýbrž ke ztrátě energie.
- 3) *Dosažení cíle v co nejkratším čase* – pro mnoho sportů (například plavání a cyklistika) je dosažení co nejkratšího času v provedení primárním cílem. U jiných sportů, například sportovních her, to ale neplatí. Například přihrávky na spoluhráče musí být správně načasované, nejrychlejší provedení není vždy to nejlepší.

Belej (2001) charakterizuje pohybovou dovednost jako způsob vykonávat pohybovou činnost správně, úsporně, vhodným způsobem a to i při změněných podmínkách. Definice Linharta (Linhart 1982 v Měkota a Cuberek, 2007, 9), Schnabela a Thiese (Schnabel a Thies 1993 v Měkota a Cuberek, 2007, 10) a Schmidta (Schmidt a Wrisberg, 2008) podle Měkoty a Cuberka (Měkota a Cuberek, 2007) „tendují k pojetí, v němž pohybová dovednost je považována za předpoklad činnosti, nikoliv za činnost samotnou“. Na druhé straně uvádějí anglicky psanou literaturu (např. Burton a Miller, 1998), ve které se za dovednost (*skill*) považuje i činnost samotná. Měkota a Cuberek (2007) dále uvádějí, že za dovednost nelze považovat každý pohyb nebo pohybovou činnost, ale důležité je cílové zaměření, tj. že dovednost implikuje určitý cíl.

Zvláštním případem pohybových dovedností jsou sportovní dovednosti. Dovalil a kol. (2008) definuje sportovní dovednosti jako „učení osvojené, v tréninku průběžně zdokonalované komponenty sportovních činností, v nichž se v integrované podobě projevují všechny podstatné faktory daného sportovního výkonu.“ Sportovní dovednost se liší od pohybové dovednosti zdůrazněním výkonové motivace a pravidly, která vymezují pohybový úkol v dané sportovní disciplíně. Schnabel a kol. (Schnabel a kol. 2003 v Měkota a Cuberek, 2007, 19) definuje sportovní dovednost jako „specifický koordinační výkonový předpoklad pro realizaci určité sportovní činnosti žádoucí technikou, který zpravidla musí být osvojen v delším učebním a tréninkovém procesu a je spojen s vytvořením regulačních mechanismů.“

Sportovní dovednosti, ať už mezi jednotlivými sporty nebo v rámci jednoho sportu, se od sebe odlišují – v délce trvání, rytmu, přesnosti pohybu, účasti smyslů, nárocích na provedení, ve způsobech efektivního nácviku apod. Z tohoto hlediska je důležité dovednosti klasifikovat a nacházet tak mezi nimi odlišnosti. Níže uvádíme klasifikaci pohybových dovedností podle Rychteckého (2002) s příklady sportovních dovedností z florbalu:

1. *Pohybové dovednosti percepční nebo motorické* (podle kritéria účasti smyslů, nervosvalového systému a výsledku) – ve florbalu je většina dovedností náročná jak na percepční aktivitu, tak na motorický projev. Podle vyjádření dominující tendence v realizaci pohybové dovednosti můžeme některé florbalové dovednosti zařadit spíše do percepčních dovedností (například přihrávka na spoluhráče, která nevyžaduje vysokou technickou náročnost, ale je důležité její načasování v kontextu řešení herní situace) nebo do motorických dovedností (například technika střelby golfovým úderem, která je především náročná na motorický projev).
2. *Pohybové dovednosti výstupově dominantní, vstupově dominantní nebo kognitivní* (podle kritéria dominance vstupu nebo výstupu) – ve florbalu jako v ostatních sportovních hrách dominují dovednosti kognitivní (herní myšlení, rozhodovací procesy a další kognitivní komponenty jsou jedním ze základních kamenů herního projevu hráče).
3. *Pohybové dovednosti diskrétní, sériové nebo kontinuální* (podle kritéria časového vztahu podnětu a reakce ve struktuře pohybu) – florbalové dovednosti řadíme jak mezi diskrétní dovednosti (s krátkým trváním a odlišením začátku a konce svého průběhu), tak mezi sériové dovednosti (s odlišením začátku a konce a zahrnující kombinace různých reakcí včetně například změny rytmu). Mezi diskrétní florbalové dovednosti můžeme zařadit všechny druhy střelby, přihrávek a zpracování míčku apod. Sériovými

florbalovými dovednostmi jsou například herní činnosti jednotlivce jako uvolňování hráče s míčkem nebo bez míčku nebo také obsazování hráče s míčkem nebo bez míčku.

4. *Pohybové dovednosti uzavřené nebo otevřené* (podle kritéria jejich vztahu k podnětu a prostředí) – všechny florbalové dovednosti vzhledem k charakteru sportovní hry řadíme mezi otevřené dovednosti z důvodu vnějšího proměnlivého prostředí ve hře s obtížně předvídatelnými změnami. Hráč se musí časově a prostorově přizpůsobovat nejen pohybu míčku, ale také činnostem svých spoluhráčů a protihráčů. Pouze při nácviku florbalových dovedností se většinou postupuje od nácviku částečně uzavřených dovedností, které se postupem času stávají otevřenými (například nácvik střelby z místa, přes nácvik z pohybu a nácvik s přítomností obránce nebo obránců až po nácvik v herních podmínkách).
5. *Pohybové dovednosti hrubé nebo jemné* (podle kritéria velikosti provedených pohybů a přesnosti pohybů) – hrubé dovednosti, pro které je charakteristické provedení velkými svalovými skupinami, jsou ve florbalu zastoupeny dovednostmi souvisejícími s pohybem hráčů na hřišti – starty nebo změny směru pohybu. Mezi jemné dovednosti zahrnující pohyby malých svalových skupin s vysokými nároky na přesnost provedení řadíme u hráčů dovednosti s florbalovou holí.

2.2.2 Motorické učení

Motorické učení zahrnuje oblast činností od základních pohybových dovedností (například lezení a chůze), kterým se člověk učí postupně od narození, přes komunikační dovednosti (například psaní a čtení) a odborné dovednosti (podle jednotlivých povolání) až k uměleckým a rekreačním dovednostem, mezi které patří také tělovýchovné a sportovní dovednosti. Nejedná se však jen o izolovaný proces učení se novým dovednostem, ale o podíl motorického učení (osvojování pohybů příčně pruhovaného svalstva) a sensorického učení (aktivní kognitivní a intelektové procesy) – znamenající psychickou aktivitu jedince při řízení pohybů. Z tohoto důvodu mluvíme o senzomotorickém učení, zahrnujícím výše uvedený psychomotorický akt. V literatuře užívaný termín „motorické učení“ je proto chápán v širším rozsahu – zahrnující i složky poznávací (kognitivní), citovou (emotivní) a snahovou (volní). (Rychtecký, 2002)

Pro trenéry, kteří se snaží učit sportovní dovednosti své svěřence, je důležité před začátkem procesu si položit tři základní otázky: „Kdo?“, „Co?“ a „Kde?“

„Kdo?“ – kdo je jedinec, kterého trénuje;

„Co?“ – jakou dovednost budu učit;

„Kde?“ – jaké mám pro učení podmínky. (Schmidt a Wrisberg, 2008)

Otázka „Kdo?“ znamená pro trenéra uvědomění si, zda se jedná o dítě nebo dospělého, začátečníka nebo pokročilého, a jestliže se s danou dovedností jedinec již setkal, tak v jaké fázi učení se nachází. Identifikace úrovně dovednosti související s fází učení je stěžejní pro zvolení efektivního tréninku vedoucího k zautomatizování dovednosti.

Popis jednotlivých fází motorického učení uvádíme podle Dovalila a kol. (2008) a Rychteckého (2002):

1. *Fáze – generalizace (iradiace)* – fáze seznámení se s dovedností a prvními pokusy. Fázi charakterizuje vysoká mentální aktivita s cílem tvorby motorického programu. Jedná se o nekoordinovaný pohyb regulovaný exteroceptory s nadměrnou aktivací svalů i aktivací svalů, které s pohybem přímo nesouvisí. Procesy zasahují různé oblasti mozkové kůry (difúzní iradiace) a způsobují generalizaci ve vnějším pohybu. Charakteristická je nízká úroveň prováděných dovedností.
2. *Fáze – diferenciací (koncentrace)* – fáze nácviku a opakování, ve které dochází ke zpevnování pohybů vedoucích ke správnému zvládnutí dovednosti. V centrální nervové soustavě dochází ke koncentraci vzruchů v souvisejících korových oblastech. Fáze je ale charakteristická neekonomičností pohybu, nekoordinovaností a křečovitostí, střední mentální aktivitou a střední úrovní prováděných dovedností.
3. *Fáze – automatizace (stabilizace)* – fáze je charakteristická ekonomickým, plynulým a koordinovaným pohybem bez soustředěné pozornosti jedince (vnitřní regulační okruh nahrazuje regulaci pomocí zrakového analyzátoru a ostatních exteroceptorů). Pro tuto fázi je charakteristický vysoký stupeň retence, nízká mentální aktivita a vysoká úroveň dovedností.
4. *Fáze – tvořivá koordinace (tvořivá asociace)* – fáze, která je uváděna především u otevřených dovedností, u které po zautomatizování pohybu jsou pro dosahování cíle nezbytné nové způsoby provedení podpořené především tvořivostí a anticipačními

schopnostmi jedince. Fázi charakterizuje vysoká mentální aktivita a mistrovská úroveň provedení dovedností.

Rychtecký (2002) rozděluje motorické učení, respektive senzomotorické učení, podle dominance kognitivních procesů, interakčních vztahů a činností učících se i učících subjektů na následující oblasti, které jsme doplnili o příklady z florbalu:

1. *Imitační učení* – představa pohybu je vytvářena přes zrakový analyzátor, dovednost je nacvičovaná jako jeden celek a mnohonásobným opakováním dochází k osvojení dovednosti (využívá se například u nácviku vedení míčku u dětí mladšího školního věku).
2. *Instrukční učení* – představa dovednosti se vytváří podle slovních instrukcí, které obsahují popis techniky nebo návod, jak při osvojování dovednosti postupovat (využívá se například u nácviku střelby u dětí staršího školního věku, většinou v kombinaci s imitačním učením).
3. *Zpětnovazební učení* – učení je založeno na principu pokusu a omylu, kdy podle úspěšnosti nebo neúspěšnosti v průběhu nácviku se provádí korekce pohybu (využívá se při nácviku herních činností jednotlivce u starších věkových kategorií, například s využitím videotechniky).
4. *Problémové učení* – je založeno na myšlenkové analýze určité problémové situace, která vyústí ve formulování hypotézy. Hypotéza je prakticky ověřována a podle výsledku buď přijata anebo zamítnuta (využívá se například u řešení situací přečíslení 2-1 nebo 3-2 z hlediska obranných činností u starších věkových kategorií).
5. *Ideomotorické učení* – je založeno na centrálním podráždění kinestetických buněk představou o pohybu nebo promyšlením pohybu. Ideomotorické učení může vhodně doplňovat praktický nácvik (může se jednat o představy herních činností jednotlivce, představy taktických řešení herních situací nebo promyšlení zásahů proti konkrétním hráčům u brankáře).

Dalšími druhy učení, které se objevují ve vztahu k motorickému učení, je explicitní a implicitní učení, které souvisí s explicitní a implicitní pamětí, respektive s vědomou a nevědomou pamětí (Atkinson 2003, Klenerová a Hynie 2010, Koukolík 2008).

Explicitní učení je takové učení, při kterém si uvědomujeme, že se učíme (Klenerová a Hynie, 2010) a souvisí s ním fáze motorického učení. Při vědomém učení se používají k osvojení dovedností instrukce (například „pokrčit kolena“, „špičky směřují vpřed“), které sportovce vedou při provádění pohybu k jeho správnému výslednému provedení. Tyto instrukce jsou uloženy v pracovní krátkodobé paměti a slouží jako podpora předváděného pohybu. Úspěšnost činnosti je hodnocena zpětnou vazbou – úspěšné pokusy mají tendenci být opakovány a uloženy v dlouhodobé paměti a neúspěšné pokusy přehlíženy. Sportovec procvičuje dovednost do té doby, než se pohyb stane automatický, nezávislý na instrukcích vytvořených v počáteční fázi učení (Masters a Maxwell, 2004).

Při nevědomém, implicitním učení, je naopak snahou neosvojovat si vědomé instrukce a naopak podle definice implicitního učení od psychologů Berryho a Dienes (Berry a Dienes 1993 v Cleeremans, 2002, 492–493), osvojovat si nové informace, aniž bychom to měli v úmyslu, a takovým způsobem, že výslednou znalost je těžké vyjádřit. V motorickém učení to znamená proces osvojování motorických dovedností bez explicitních instrukcí, který by měl vést k předcházení selhání při sportovním výkonu pod vlivem stresu (Masters, 1992) a který je založen na praktických postupech jako je učení s druhým úkolem, učení s redukcí chyb, analogickém učení nebo učení využívající „vnějšího ohniska“ pozornosti (Dragounová a kol. 2013, Liao a Masters 2001, Masters a kol. 2004, Maxwell 2008, Maxwell a kol. 2000, 2001, Poolton 2007).

2.3 Mladší školní věk a jeho charakteristika

Období mladšího školního věku můžeme rozdělit na dvě období, přibližně od 6 do 9 let a od 9 do 11 až 12 let (Choutková a Kučera 1970, Kirchner 1981, Perič 2012). První období označují autoři za dětství a druhé období za prepubescenci (Perič, 2012) nebo za mladší školní věk nižšího stupně a střední stupeň mladšího školního věku (Choutková a Kučera, 1970). V propracovaném kanadském systému dlouhodobého vývoje sportovce „*Long term athlete development*“ (Balyi a kol. 2005, Balyi a Ross 2009) mají období mladšího školního věku zařazeno od 6 let s neomezenou horní věkovou hranicí v závislosti na individuálním nástupu růstového spurtu.

První období mladšího školního věku je charakterizováno progresivními změnami v oblasti pohybové, intelektuální a volní (Choutková a Kučera, 1970).

V pohybové oblasti dochází ke změnám pákových poměrů končetin a trupu a nastávají příznivé podmínky pro vývoj různých pohybových forem (Perič, 2012). Výška a váha dětí vzrůstá rovnoměrně, přibližně o 5 cm a 2,7 kg ročně, srdce a plíce ještě nejsou plně vyvinuty a tepová a dechová frekvence ukazují postupný pokles. Okolo 9 let je zřídka tepová frekvence nad 90 tepů za minutu (Kirchner, 1981). Vývojové zvláštnosti tak nevytvářejí vhodné podmínky pro soustředěnější vytrvalostní a silový rozvoj (Štilec, 1989). Naopak nervový systém je dostatečně zralý pro složitější koordinační pohyby a učení se novým pohybům, jedná se o příznivé období pro vznik nových podmíněných reflexů (Perič, 2012). Začátkem školního věku, kdy motorika dosahuje relativně vysokého stupně vývoje, jsou pohyby dětí poměrně plynulé a vyvážené a nové dovednosti jsou poměrně lehce a rychle zvládnuty, ale při nedostatečném opakování rychle zapomínány (Štilec, 1989). Charakteristickými znaky tohoto období jsou snadná únava a rychlé zotavení. Koordinace oko–ruka ještě není zcela vyvinuta. Reakční čas je pomalý, ale ukazuje trvalé navyšování během tohoto období (Kirchner, 1981).

V intelektuální oblasti dítě cvičí paměť a dokáže ji využívat. Je možné i trénovat pozornost. Děti mohou být náladové a popudlivé, charakterizuje je impulsivnost, přechody z radosti do smutku a naopak. Jednostrannou práci se rychle unavují. Slabě je zatím vyvinuta vůle, nedokáže dlouhodobě sledovat cíl a soustředit se na jednu činnost. Myšlení je konkrétní, a proto je třeba názorných příkladů (Choutková a Kučera 1970, Štilec 1989). Děti jsou vnímavé k okolnímu prostředí a faktorům, které odvádějí pozornost. Důležitá je krátká doba, po kterou se děti dokáží koncentrovat (Perič, 2012). Děti jsou extrémně kreativní, rády

opakují aktivity, které znají a umějí. Mají rády dobrodružství a nebojí se (Kirchner, 1981). Zvláště u chlapců byl pozorován patrný nárůst odvahy (Štílec, 1989).

Z důvodu přechodu od hravého dětství k vážné práci ve školní lavici, dochází ke zvýšení počtu hodin sezením bez možnosti vyrovnání relativně nové polohy těla (Choutková a Kučera, 1970). Správné držení těla je zásadním problémem právě v druhém, předpubertálním, období mladšího školního věku (Kirchner, 1981), kdy je třeba věnovat častou pozornost návyku dobrého držení těla (Štílec, 1989). V tomto období odpovídá velikost a kapacita srdce a plic přiměřeně nárůstům výšky a váhy. Výška a váha dětí vzrůstá přibližně o 5 – 6 cm a 3 kg ročně (Balyi a Ross 2009, Kirchner 1981). Tepová frekvence se pohybuje kolem 12. roku od 80 do 90 tepů za minutu. V tomto období, oproti předchozímu, je možné zařazovat již delší vytrvalostní úseky. Snižuje se ale úroveň pohyblivosti, především u chlapců. Zlepšuje se svalová koordinace, koordinace oko–ruka a snižuje se reakční čas (Kirchner, 1981). Období od 8 do 10 let se často nazývá „zlatým věkem motoriky“, pro který je charakteristické rychlé učení se novým pohybům a kvalita provedení pohybu. Děti jsou schopné provádět i koordinačně náročná cvičení, která mohou být ještě z počátku období mladšího školního věku pro děti náročná (Perič, 2012).

U dětí se zvyšuje intelektuální zvědavost a kontrola emocí (Kirchner, 1981). Děti prožívají postupné období socializace, ve kterém se začleňují do kolektivu a přizpůsobují se daným zákonitostem a pravidlům. Na konci období mladšího školního věku, vzhledem k prepubertální fázi, nastává fáze kritičnosti v hodnocení jevů a podnětů ze sociálního prostředí (Perič, 2012).

Desetileté děti mají už zájem o výsledky pohybu, ale porovnání s druhými dětmi (vítězství v utkání, poměr branek) je přitažlivější dříve než srovnávání vlastních výkonů. Motiv soutěživosti a potřeba sebeuplatnění jsou v tomto věku dostatečně rozvinuty (Choutková a Kučera, 1970).

2.4 Diagnostika pohybových dovedností

Diagnostika pohybových dovedností se soustřeďuje na stupeň osvojení určité pohybové dovednosti, případně na šíři uplatnění pohybové dovednosti, tj. míru jejího zobecnění. Příkladem může být diagnostikování plavecké dovednosti – první parametr diagnostikujeme změřením uplavané vzdálenosti a druhý parametr naopak zjištěním, že testovaná osoba umí plavat prsa i znak (Měkota a Cuberek, 2007).

V diagnostice pohybových dovedností se uplatňují dva základní přístupy. První je založen na změření pohybového výkonu a diagnostickým nástrojem tohoto přístupu jsou motorické testy (v případě sportovních dovedností testy sportovních dovedností). Druhý přístup je založen na posouzení průběhu pohybového aktu a diagnostickým nástrojem je škálování.

Testy a baterie testů pro děti předškolního a mladšího školního věku zaměřené na testování pohybových dovedností nejčastěji slouží k určování motorického vývoje s cílem včasného odhalení motorických poruch.

Pro nejmladší děti existují také testy základních motorických dovedností (viz „Test běhu u tříletých dětí“). „Základní motorické dovednosti, jako chůze, běh a skok, jsou klasickými testovými metodami postižitelné pouze v raném věku. V dospělosti dosažený čas při běhu nebo překonaná vzdálenost při skoku budou spíše indikátory motorických schopností“ (Měkota a Blahuš, 1983, 242). Výjimkou je pouze dovednost přesného házení hodnocená testy, které zahrnují hody na cíl. Příkladem může být test házení tenisového míče na cíl určený školním dětem (Mecner, 1975).

Měkota a Cuberek (2007) označují jako tři hlavní možnosti diagnostiky dovedností sportovního charakteru *testování, posuzování (škálování) a samotný výkon ve sportovní disciplíně*. U některých sportů (například střelba a lukostřelba) lze podle autorů ztotožnit samotný výkon ve sportovní disciplíně, tj. výsledek v soutěži, s výsledkem testu příslušné dovednosti.

Mezi velmi dobře testovatelné sportovní dovednosti uvádí Měkota a Blahuš (1983) dílčí dovednosti jednotlivých míčových her. Testy navržené pro herní činnosti jednotlivce v basketbalu, volejbalu, házené, fotbalu, ale i testy tenisových, badmintonových, softbalových a dalších dovedností autoři rozdělují do šesti skupin:

- a) stěna a míč,
- b) hod (kop) míčem na cíl,
- c) lokomoce s míčem,
- d) hod (kop) míčem na vzdálenost,
- e) žonglování míčem,
- f) řetězec pohybových činností s míčem.

Mezi dobře testovatelné dovednosti dále uvádí Měkota a Blahuš (1983) plavání, u kterých je čas nutný k uplavení určité vzdálenosti nebo délka trati uplavaná v určitém čase indikátorem plavecké dovednosti. Indikátorem plavecké dovednosti je také účinnost plaveckých záběrů, kterou lze posoudit podle vzdálenosti uplavané určitým počtem temp (Měkota a Cuberek, 2007). Atletické disciplíny jsou naopak špatně testovatelné, protože jsou výkony ve většině atletických disciplín silně ovlivněny kondičními schopnostmi.

2.4.1 Kvalitativní a kvantitativní výzkum

Kvantitativně orientovaný výzkum je klasickým výzkumem vedle kterého se v posledních letech ve světě i u nás silně rozvinul i kvalitativně orientovaný výzkum. Mezi kvantitativním a kvalitativním výzkumem existuje několik důležitých rozdílů (Gavora 2010, Chráska 2007, Švec a kol. 2009):

- *odlišné cíle výzkumu* – kvantitativní výzkum je zaměřen především na vysvětlení jevu, tj. explanace příčin a predikce následků, kvalitativní výzkum je zaměřen spíše na interpretaci a porozumění významu;

- *odlišná metodologická orientace* – kvantitativní metodologie výzkumu se zařazuje mezi analytické metodologické orientace a naopak kvalitativní mezi syntetické metodologické orientace;
- *odlišný výběr* – v kvantitativním výzkumu se vybírají zkoumané osoby tak, aby reprezentovali určitou populaci, kvalitativního výzkumníka zajímá konkrétní případ;
- *odlišná interpretace výsledků* – statistická sumarizace výzkumných výsledků v kvantitativním výzkumu a narativní shrnutí výsledků kvalitativních výzkumů;
- *odlišný postoj výzkumníka* – výzkumník v kvantitativním výzkumu se snaží držet odstup od zkoumaných jevů, naopak výzkumník v kvalitativním výstupu se snaží o sblížení se zkoumanými osobami.

Hendl (2015) uvádí jako přednosti kvantitativního výzkumu testování a validizaci teorií, možnost zobecnění výsledků na populaci, možnost eliminace působení rušivých proměnných, přesná numerická data a jejich relativně rychlý a přímočarý sběr a analýzu, relativně nezávislé výsledky na výzkumníkovi a možnost zkoumání velkých skupin. Naopak jako nevýhody uvádí, že kategorie a teorie použité výzkumníkem nemusejí odpovídat lokálním zvláštnostem, získaná znalost může být příliš abstraktní a obecná pro přímou aplikaci v místních podmínkách a výzkumník je omezen reduktivním způsobem získávání dat. Mezi výhody kvalitativního výzkumu Hendl (2015) řadí možnost studovat procesy a navrhnout teorie, zkoumat fenomén v přirozeném prostředí, získávat podrobný popis a vhléd při zkoumání osob nebo jevů, dobře reagovat na místní situace a podmínky, hledat lokální příčinné souvislosti a pomáhat při počáteční exploraci fenoménů. Za nevýhody naopak považuje obtížné provádění kvantitativní predikce, obtížnější testování hypotézy a teorie, časově náročnou etapu sběru a analýzy dat, možnost ovlivnění výsledků výzkumníkem a jeho osobními preferencemi a fakt, že výsledná znalost nemusí být zobecnitelná na populaci a do jiného prostředí.

Kvantitativní a kvalitativní výzkum využívá v současné době poměrně velké množství metod, které se rozšířili především s rozvojem a dostupností technologií (například videostudie, nepřímé pozorování nebo počítačová analýza textu).

Mezi metody kvantitativně orientovaného výzkumu patří například pozorování, dotazník, interview, experiment a odborné posuzování. Odborné posuzování využívá techniky kategoriálních systémů nebo škálování (Gavora 2010, Jansa 2012, Švec a kol. 2009).

2.4.2 Škálování

Škálování se zařazuje k technikám odborného posuzování patřící mezi kvantitativní empirické metody. Uplatňuje se při diagnostice pohybových dovedností založené na pozorování, registraci a zhodnocení průběhu pohybového aktu. Ke škálování používáme posuzovací škály (stupnice). Výsledky pozorování se umisťují na numerickou nebo grafickou škálu – tzv. kontinuum, kolem kterého se umisťují posudky. Často užívaným typem škály je posuzovací škála kumulativní, rozdělená na dílčí činnosti posuzované samostatně, a kombinovaná škála, která postihuje větší počet hledisek posuzovaného jevu. (Měkota a Blahuš, 1983)

Obecně je možné posuzovací škály charakterizovat jako nástroj, který zjišťuje míru vlastnosti jevu nebo jeho intenzitu a posuzovatel vyjadřuje své hodnocení určením polohy na škále (Gavora, 2010). Posuzovací škály pomáhají objektivizovat subjektivní hodnocení vytvořením definovaného systému skórování (Měkota a Cuberek, 2007).

Škálováním se v Čechách zabýval psycholog Břicháček (1978). Pro psychologické škálování uvádí několik škálovacích technik: metoda párového srovnávání podnětů, metoda pořadových stupnic, technika zdánlivě stejných intervalů, škálogramová analýza a další. Výše uvedené techniky se staly základem také pro tvorbu motorických škál určených k testování pohybových dovedností. Například techniku zdánlivě stejných intervalů použila ve své kinantropologické studii Kavalířová (2003) pro posuzování pohybových dovedností ve sportovní gymnastice při výuce tělesné výchovy a Čepička (2003, 2005) použil škálogramovou analýzu k posuzování úrovně motorické dovednosti při manipulaci s míčem.

Škálování vychází z teorie měření, ale o škálování jako o měření můžeme hovořit pouze jako o měření v širším smyslu nebo na nižší úrovni (Berka, 1977). U škálování se využívá ordinální stupnice splňující topologizační podmínky metrizace a aditivity (tj. čítání). Pomocí topologizace se zjišťuje kvazi-uspořádání nebo kvazi-řada kvalitativního pojmu. Topologizačními podmínkami rozumíme existenci tranzitivního vztahu „rovno“ a „menší než norma“ (tzv. tranzitivita) a rozhodnutí zda platí vztah „rovno“ nebo „menší než norma“ (tzv. rozhodnutelnost). Škálování je metodou kvazi kvantifikační, protože splňuje pouze tři podmínky metrizace (tranzitivitu, rozhodnutelnost a aditivitu). Čtvrtou podmínku metrizace, konstantnost jednotky (neměnné velikosti intervalu), splňují pouze fyzikální veličiny (Blahuš 1996, Jansa a kol. 2012).

Ke škálování používáme posuzovací škály (stupnice). U posuzovacích škál je na rozdíl od měření kvantifikace kvalitativních znaků v některých případech zjednodušená a škály mají charakter pouze pořadových stupnic (Jansa a kol., 2012). Obecně je možné posuzovací škály charakterizovat jako nástroj, který zjišťuje míru vlastnosti jevu nebo jeho intenzitu a posuzovatel vyjadřuje své hodnocení určením polohy na škále (Gavora, 2010).

Škály se dělí podle množství sledovaných jevů (znaků, proměnných) na škály unidimenzionální a multidimenzionální.

Obecným dělením škál je pak rozdělení na škály materiální a škály konceptuální. Konceptuální škály jsou charakterizované určitým uspořádáním škálových (numerických) hodnot, které lze přiřazovat veličinám. Konceptuální škála je vymezena charakterem numerických hodnot, počátkem a povahou distance mezi dvěma libovolnými sousedními škálovými hodnotami (Berka 1977, Měkota a Blahuš 1983).

Druhy konceptuálních škál (Měkota a Blahuš 1983, Chráska 2007, Švec a kol. 2009) uvádíme s podmínkami měřitelnosti podle Berky (Berka 1972 v Měkota a Blahuš, 1983):

- *nominální* – třídění neboli klasifikace objektů do kategorií podle kvalitativních znaků, například třídění žáků podle pohlaví nebo zdravotního stavu (podmínky měřitelnosti: empirické relaci „stejně velký/silný...“ odpovídá aritmetická relace „=“; statistické operace: procentuální vyjádření četnosti, modus, chí–kvadrát...),
- *ordinální (pořadová)* – určování pořadí objektů (vzestupné nebo sestupné) například podle významnosti nebo oblíbenosti; kvantitativní rozdíly mezi subjekty jdoucími za sebou nemusí být shodné (podmínky měřitelnosti: empirické relaci „starší/větší/silnější...než“ odpovídá aritmetická relace „>“ a empirické relaci „stejně velký/silný...“ odpovídá aritmetická relace „=“; statistické operace: medián, Spearmanův a Kendellův koeficient...),
- *intervalová* – měření kardinálních znaků (např. škála na hodnocení vlastností žáka), nejčastěji se stejnými intervaly mezi hodnotami (podmínky měřitelnosti: využití aritmetické relace „+“ a u stejně velkých intervalů mezi hodnotami i součinu; intervalové škály nemají tzv. absolutní nulovou hodnotu, ale stanovuje se centrální hodnota; statistické operace: aritmetický průměr, směrodatná odchylka, párový t-test, analýza rozptylu, Pearsonův koeficient...),

- *poměrová (podílová)* – měření např. výkonů nebo zpracování didaktických testů; nejčastěji jsou přiřazována stupnicím přirozená čísla umožňující jakékoliv matematické a statistické operace (podmínky měřitelnosti: viz intervalová stupnice).

Při používání posuzovacích škál Jansa a kol. (2012) a Gavora (2010) upozorňují na několik druhů zkreslení: nadhodnocování (zkreslení z důvodu mírnosti), podhodnocování (zkreslení z důvodu přísnosti), tzv. centrální tendence (vyhýbání se krajnímu hodnocení), tzv. haló efekt (hodnocení dané vlastnosti pod vlivem jiné vlastnosti) a další.

2.4.3 Definice a vlastnosti motorického testu

Motorický test je diagnostická metoda určená k zjišťování úrovně motorických předpokladů. „Motorickým testem rozumíme standardizovaný postup (zkoušku), jehož obsahem je pohybová činnost a výsledkem číselné vyjádření průběhu či výsledku této činnosti“ (Čelikovský a kol., 1979). Mezi základní vlastnosti motorického testu patří objektivita, validita a reliabilita.

Objektivita se týká kvality výsledků motorického testu a je určena stupněm shody testových výsledků, které získají současně různí examinátoři (Měkota a Blahuš, 1983). Vysoká hodnota koeficientu objektivit zaručuje reprodukovatelnost testu – jedná se o stupeň nezávislosti výsledků na výzkumníkovi nebo testované osobě.

Validita je považována za nejdůležitější vlastnost motorického testu a znamená platnost motorického testu – schopnost motorického testu zjišťovat to, co má být motorickým testem zjišťováno. V teorii a praxi rozlišujeme několik druhů validity:

- a) *Obsahová validita* – zjišťovat obsahovou validitu testu znamená „hodnotit adekvátnost pohybového obsahu testu a posuzovat vhodnost výběru položek nebo subtestů s ohledem na účel testování“ (Měkota a Blahuš, 1983). Obsah testů by měl být reprezentativním výběrem například osvojených dovedností určitého sportu (test k měření tenisové dovednosti nebude obsahovat jen položku z tenisového podání, ale musí obsahovat i forhendový a bekhendový úder, volej apod.).

- b) *Kriteriální validita* – jedná se o míru shody mezi výsledky testu a výsledky jiného měření podle ověřeného kritéria. Rozlišujeme dva druhy kriteriální validity:
- souběžná validita – zjišťuje se srovnáním výsledků jednoho testu s výsledky jiného testu, který je validní a je zadán shodným subjektem ve stejném čase,
 - predikční validita – je shodná se souběžnou validitou s tím rozdílem, že srovnání s výsledkem jiného testu se uskutečňuje s časovým odstupem.
- c) *Konstruktová (teoretická) validita* – stanovuje, jestli test zjišťuje právě ten konstrukt (například dovednost), pro jehož měření byl sestrojen.

Reliabilita znamená přesnost a spolehlivost výzkumného nástroje a vyjadřuje velikost chyb testování. V praxi odhadujeme koeficient reliability motorického testu pomocí jeho paralelní formy za použití korelace. Pro vytvoření paralelní formy testu se používají čtyři postupy (Měkota a Blahuš, 1983):

- a) *opakování měření* daného testu (metoda stability – opakované měření stejným testem),
- b) *dělení daného testu* (metoda konzistence – po rozpůlení testu, nebo v některých případech rozdělení na více částí, se zjistí stupeň shody ve výsledcích obou polovin, nebo všech částí, pomocí korelačního koeficientu),
- c) *testování jiným paralelním testem* (metoda ekvivalence – použití dvou ekvivalentních testů, výsledky obou měření se srovnávají pomocí korelačního koeficientu),
- d) *analýza rozptylu* (používá se při více než jedné paralelní formě).

Standardizovaný motorický test by měl mít také vypracovaný systém skórování a hodnocení testových skóre pomocí výkonnostních norem. Norma je obvykle kvantitativní hodnota, empiricky určená, představující normální, tj. obvyklý, výkon u normové (odpovídající) populace. Normy se odvozují z výsledků rozsáhlých a reprezentativních šetření (Měkota a Blahuš, 1983).

Pokud je několik testů standardizováno společně a jejich výsledky se kumulují a vytvářejí tak jeden souhrnný výsledek, jedná se o testovou baterii.

Obdobné rozdělení a charakteristiky jednotlivých vlastností standardizovaných testů uvádí také autoři Gavora (2010) a Švec a kol. (2009) v oblasti pedagogického výzkumu.

2.4.4 Testy a testové baterie

Pro účely naší práce jsme provedli rozsáhlé šetření po dostupných testech pohybových a sportovních dovedností pro děti mladšího školního věku. Výsledkem bylo zjištění, že oproti nalezenému velkému množství testů a testových baterií určených k zjišťování úrovně pohybových schopností, je oblast dovednostních testů výrazně chudší. Poměrně rozsáhlá byla pouze oblast motorických testů určených k diagnostice motorických poruch u dětí. Velké množství motorických testů pochází z amerického kontinentu (Burton a Miller, 1998). Jedná se o testy určené dětem od narození do konce mladšího školního věku nebo o testy pro děti pohybově nebo mentálně handikepované. Zajímavé je, že většina testů pro zjišťování úrovně pohybových dovedností vznikla v průběhu minulého století. Domníváme se, že se vědecká pozornost v novém tisíciletí přesunula především do neprobádané oblasti hendikepovaných nebo do oblastí rehabilitační (například Van de Winckel a kol. 2006, Velozo 2008).

Nejsložitější situace byla u vyhledávání testů sportovních dovedností. Testy sportovních dovedností jsou většinou určeny až starším dětem a dospělým, přestože senzitivní období pro učení se novým dovednostem u jednotlivých sportů se vztahuje k mladšímu školnímu věku (Perič, 2012).

V Čechách vzniklo několik testů sportovních dovedností v sedmdesátých a osmdesátých letech minulého století, v současné době se jedná jen o výjimečné pokusy. V zahraničí jsou nejlépe propracované testy zkonstruované v šedesátých letech asociací AAHPERD pro několik vybraných sportů (na americkém kontinentu rozšířených) a přepracované v osmdesátých a devadesátých letech minulého století. Snaha asociace vyústila v kvalitní nástroj určený pro testování dovedností začínajících hráčů. Dovednostní testy také mají zkonstruované některé sportovní svazy nebo unie nebo se objevují na webových stránkách sportovních organizací nebo oddílů, u testů však chybí informace o vědecké standardizaci, pokud byla udělána. Jako příklad uvádíme sešit autorů Hrazdíry a Maršálka (2002) nazvaný *Celoplošné testování hráčů pozemního hokeje všech věkových kategorií*, zkonstruovaný bez vědecké standardizace (Tůma, 2004).

Ve florbalu jsme objevili jen baterii testů pro děti od švédské florbalové unie (Karlberg, 2013), která ale postrádá jakýkoliv záznam o vědecké standardizaci a dále standardizovaný test pro dospělé hráče v diplomové práci od Tůmy (2004), který výše uvedené testy z pozemního hokeje převedl a standardizoval na test florbalových dovedností.

Následující přehled standardizovaných testů, určených pro děti mladšího školního věku nebo s možností testování i této věkové kategorie, je rozdělen na testy pohybových a sportovních dovedností zkonstruované v českých zemích a v zahraničí. Testy a baterie testů jsou uspořádány abecedně podle názvu.

Výběr testů pohybových dovedností zkonstruovaných v ČR:

- **Test házení tenisového míče na cíl** (Mecner 1975 v Měkota a Blahuš 1983, 247) – test určený školním dětem se skládá z hodů tenisovým míčkem ve stoji, preferovanou rukou, na terč vzdálený 10 m.
- **Test úrovně somatického a motorického vývoje předškolních dětí** (Berdychová a Pařízková 1975 v Měkota a Blahuš 1983, 264) – test určený pro děti od 3 do 6 let složený z antropometrických ukazatelů a sestavy pěti motorických položek: skok daleký z místa, hod kriketovým míčkem, apod.
- **Test běhu u tříletých dětí** (Kučera 1985 v Měkota a Cuberek 2007, 31) – test běhu na 10 m letným startem zjišťuje, zda tříleté děti ovládají dovednost běhu.

Výběr testů sportovních dovedností zkonstruovaných v ČR:

- **Posuzovací škála – gymnastika** (Kavalířová, 2003) – pětibodové posuzovací škály zkonstruované pro potřeby učitelů tělesné výchovy a určené k posuzování základních dovedností ze sportovní gymnastiky – přemetu stranou a roznožky přes kozu.
- **Posuzovací škála – gymnastika** (Gajda 1981 v Měkota a kol. 1988) – kombinovaná škála (grafická a numerická) určená pro hodnocení dovednosti kotoul vzad do zášvihu.
- **Posuzovací škála – házená – vrchní přihrávka jednoruč** (Šafaříková 1972 v Měkota a Blahuš 1983, 240) – vertikální numerická a grafická škála je určená pro kvantifikaci dovednosti přihrávat v házené.
- **Posuzovací škála – plavání** (Řehoř 1969 v Měkota, 1973) – sestava dvanácti kombinovaných škál (numerických a deskriptivních) určená pro posouzení některých plaveckých činností – dýchání, splývání, plavecké styly (prsa, kraul, znak), skoky do vody, potápění.

- **Posuzovací škála – volejbal – předpoklad pro manipulaci s míčem** (Čepička, 2005) – posuzovací škála určená k diagnostice předpokladu pro manipulaci s míčem. Diagnostický nástroj slouží k výběru dětí do sportovních tříd základních škol se zaměřením na volejbal.

Výběr testů pohybových dovedností zkonstruovaných v zahraničí:

- **APEAS – „Los Angeles Unified School District Adapted Physical Education Assessment Scale“** (Seaman a DePauw 1989 v Burton a Miller 1998, 343) – test je určený dětem od 5 do 12 let s cílem správného zařazení dětí do tělesné výchovy. Test se skládá z 18 testových položek obsahující například testy hodnotící dovednosti s míčem nebo lokomoci.
- **Gutteridge Rating Scale of Motor Skill** (Gutteridge 1939 v Burton a Miller 1998, 342) – kvalitativní posuzovací škála o 14 úrovních aplikovaná na základní motorické dovednosti (skákání, házení, chytání, driblování apod.). Škála je určena pro děti od 2 do 6 let a hodnocení je rozdělené podle věku a pohlaví.
- **KTK – „Körper – koordination Test für Kinder“** (Kiphard a Schilling 1974 v Měkota a Blahuš 1983, 266) – test určený pro děti od 5 do 14 let. Cílem testu bylo odhalit pohybovou (koordinační) „zaostalost“ dětí. Test se skládá ze čtyř položek: chůze vzad po kladinách, skoky snožmo přes lištu, přeskoky destiček jednož, chůze po deskách překládaných z jedné strany na druhou.
- **MABC Test – „Movement Assessment Battery for Children Test“** (Henderson a Sugden 1992 v Burton a Miller 1998, 346) – test je určený pro děti od 4 do 12 let. Jeho cílem je identifikovat jedince s pohybovými poruchami a dohromady s „MABC Checklist“ plánovat možné intervence. Test se skládá z 32 položek zahrnujících testování manuálních dovedností (například navlékání korálek), dovedností s míčem (například chytání, házení, kutálení) a pohybových schopností (statická i dynamická rovnováha).
- **MOT 4 – 6** (Zimmer 1979 v Adamírová a kol., 2010, 20) – test určený pro děti od 4 do 6 let. Test se skládá z osmnácti položek, které zahrnují obratnost a koordinaci, jemnou motoriku, rovnováhu, reakční schopnost, sílu odrazu, hbitost a pohybovou přesnost – hod na cíl a vkládání zápalek do krabičky.

- **Ozereckého testy motorické vyspělosti** (Ozereckij 1930 v Neuman, 2003, 120) – test zkonstruován roku 1923 ruským neurologem N. J. Ozereckým byl určen k testování dětí a mládeže ve věku od 4 do 16 let. Cílem testů bylo určit motorický věk dětí. Testy se hodnotily plusovými a minusovými znaménky za úspěšné, respektive neúspěšné splnění motorického úkolu.
- **PPMS – „Purdue Perceptual – Motor Survey“** (Roach a Kephart 1966 v Burton a Miller 1998, 349) – test obsahuje 30 položek rozdělených do 11 subtestů (z motorických dovedností zahrnující například překážkovou dráhu) hodnocených na čtyřbodové škále. Položky jsou hodnoceny odděleně pro děti z prvních až čtvrtých tříd prvního stupně amerických škol. Cílem testů je identifikovat chyby v motorickém vývoji u dětí mladšího školního věku.
- **TGMD – „Test of Gross Motor Development“** (Ulrich 1979 v Burton a Miller 1998, 227 – 242) – test hrubě motorického vývoje je určený dětem od 3 do 10 let. Test diagnostikuje 12 základních dovedností rozdělených do dvou subtestů – testy lokomočních a manipulačních dovedností. Z výsledku testu lze vypočítat kvocient všeobecného motorického rozvoje GMDQ („Gross Motor Development Quotient“). Test je populární v USA a k testu také existují normy odvozené testováním 908 amerických dětí z osmi států (Měkota & Cuberek, 2007).
- **TMP – „Test of Motor Proficiency“** (Gubbay 1973 v Burton a Miller 1998, 351) – test zahrnuje osm položek testující hrubé i jemné motorické dovednosti a je určen dětem ve věku od 8 do 12 let. K dispozici jsou normy pro sedmileté děti.

Výběr testů sportovních dovedností zkonstruovaných v zahraničí:

- **AAHPERD Sport Skill Test Series** – americká asociace AAHPERD (původně AAHPER) od roku 1959 připravovala a postupně realizovala projekt, jehož cílem bylo vypracování jednotně koncipovaných testů sportovních dovedností. Testy určené pro basketbal, softbal a volejbal měřily základní dovednosti daného sportu (například u basketbalu bylo předmětem testování několik druhů střelby, ovládání míče, dribling, výskok nebo přihrávka). Testy byly určeny pro učitele (jako pomůcka pro posuzování pokroku u jednotlivců) nebo pro trenéry (pro odhalení talentů nebo k diagnostice chyb v technice). Testy obsahovaly normy ve formě procentilů (například u basketbalu získané zpracováním 800 údajů pro každý věk i pohlaví od 10 do 18 let). V osmdesátých a devadesátých letech byly testy upraveny a v roce 1989 vznikl test

tenisových dovedností. Testy jsou primárně určeny pro začínající sportovce. (Baumgartner a kol., 2003; Měkota, 1973; Měkota a Blahuš, 1983; Morrow a kol., 2000)

Test basketbalových dovedností AAHPERD (Hopkins, Shick a Plack, 1984) – původní testová sestava z roku 1966. K testu existují normy pro chlapce a dívky ve věku od 10 do 18 let. Testový profil se skládá z devíti testů, které zahrnují střelbu na koš, zpětnou přihrávku, výskok, přihrávku na přesnost a dribling.

Test softbalových dovedností AAHPERD (Rikli, 1991) – původní testová sestava z roku 1966. K testu existují normy pro chlapce a dívky ve věku od 10 do 18 let. Testový profil se skládá ze sedmi testů, které zahrnují například nadhazování, odpálení, chytání a přeběhy.

Test tenisových dovedností AAHPERD (Hensley, 1989) – testová sestava z roku 1989. Testová baterie se skládá ze tří jednotlivých testů, které zahrnují základní údery (forhend a bekhend), podání (při podání se hodnotí přesnost a razantnost) a volej od sítě.

Test volejbalových dovedností AAHPERD (1984) – testová sestava z roku 1969. K testu existují normy pro chlapce a dívky ve věku od 10 do 18 let. Testový profil se skládá ze čtyř testů, které zahrnují odbíjení o stěnu, umístěné podání, přihrávku a nahrávku.

- **Fotbalová baterie testů** (Yeagley 1972 v Baumgartner a kol. 2003, 430) – baterie testů vytvořená k měření základních fotbalových dovedností začínajících hráčů. Baterie testů je složená z testu fotbalového driblinku, testu přihrávání míče o zeď nohou (volej), testu žonglování s míčem a testu hlavičkování.
- **NCSU Volleyball Service Test** (Barlett a kol. 1991 v Morrow a kol. 2000, 308) – North Carolina State University test je určený k hodnocení volejbalového podání (podání vrchem i spodem).
- **Posuzovací škála – badminton** (Baumgartner a kol., 2003) – škála, u které se ve hře posuzují čtyři oblasti (podání, údery, strategie hry a postavení hráče na hřišti včetně práce nohou), které jsou rozdělené na čtyři podoblasti a každá je popsána deskriptivním výrokem.

- **Posuzovací škála – plavání** (Baumgartner a kol., 2003) – posuzovací škála určená k hodnocení plaveckého stylu znak. Hodnoceny jsou pohyby paží a dolních končetin, provedení tempa (například koordinace paží a dolních končetin) a také efektivita záběru.
- **Posuzovací škála – volejbal** (Baumgartner a Jackson 1975 v Měkota a Blahuš 1983, 241) – kombinovaná posuzovací škála se skládá z devíti znaků, které examinátor posuzuje 1 až 5 body. V testu se hodnotí podání (výška nad sítí, přesnost umístění nebo obtížnost vrátit míč), smeč (přesnost umístění, obtížnost vrátit míč nebo spolupráce s nahrávačem) nebo nahrávka (výška nad sítí, přesnost umístění nebo spolupráce se smečařem) a kolektivní hra (pohyb na hřišti, vybírání míčů a zpracování obtížných míčů, týmová spolupráce).
- **Test badmintonového podání – Sebolt Short Service Test** (Sebolt 1968 v Baumgartner a kol. 2003, 424) – cílem testu je měření úspěšnosti badmintonového podání – „krátkého“ servisu. Testová osoba má 20 pokusů s cílem umístit servis blízko sítě.

3. GUTTMANOVA ŠKÁLA

Guttmanovu perfektní škálu, která bývá nazývána také škálogramem, vypracoval Louis Guttman na začátku čtyřicátých let minulého století a měla sloužit k psychologickému škálování. Jeho cílem bylo vytvořit jednodimenzionální a homogenní škály, ve kterých se všechny položky vztahují k jedinému faktoru, ale zároveň se liší svou závažností a jsou řazeny od nejllehčí položky k nejtěžší.

Guttmanova perfektní škála, neboli škálogram, je řada podnětových položek sestavena takovým způsobem, aby bylo možno konstatovat, že testovaná osoba, která odpoví kladně na *x*-té položce, získá vyšší skóre než testovaná osoba, která na ni neodpoví. Znamená to také, že testovaná osoba, která odpoví na položku s určitou hodnotou, musí odpovědět pozitivně i na všechny položky, které mají nižší hodnotu. Pokud známe skóre, kterého testovaná osoba dosáhla, pak můžeme jednoznačně stanovit strukturu jejich odpovědí, tj. určit, u kterých položek byla odpověď pozitivní a u kterých položek byla odpověď negativní (Břicháček, 1978). V kinantropologickém výzkumu rozumíme pozitivní a negativní odpovědi úspěšnost a neúspěšnost testované osoby na dané položce.

Břicháček (1978) popisuje postup při vytváření unidimenzionální škály (škály měřící jednu vlastnost) Guttmanova typu ve dvou fázích:

- 1) vytvoření žebříčku položek (položky by měly vytvořit v optimálním případě monotónní lineární funkci);
- 2) sestavení intenzivní složky škálogramu (tj. nalezení nulového bodu, ve kterém se negativní postoj mění v pozitivní).

Při výběru položek se doporučuje volit větší počet položek (alespoň 12), protože „čím více položek je k dispozici, tím spíše se podaří z nich vybrat část, která splňuje požadavky škálogramu...“ (Břicháček, 1978). Vybrané položky se pak předloží většímu počtu testovaných osob a výsledky se zapíší do „škálogramové tabulky“. Břicháček (1978) uvádí postup na příkladu s 5 položkami (odstupňovanými podle obtížnosti nebo intenzity postoje), na které testované osoby odpovídají „ano“ nebo „ne“ (*tabulka č. 1*).

Tabulka č. 1 Příklad výchozích dat pro sestavení škálogramu (Břicháček, 1978)

p.o.	POLOŽKY										kladné body	chyby
	A		B		B		C		D			
	1	0	1	0	1	0	1	0	1	0		
1	x		x		x		x		x		5	0
2	x		x		x		x		x		5	0
3	x		x		x		x		x		5	0
4	x		x		x		x			x	4	0
5	x		x		x		x			x	4	0
6	x		x		x			x	x		4	2
7	x		x		x			x	x		4	2
8	x		x		x			x		x	3	0
9	x		x		x			x		x	3	0
10	x		x		x			x		x	3	0
11	x		x			x	x			x	3	2
12	x		x			x	x			x	3	2
13	x		x			x		x		x	2	0
14	x		x			x		x		x	2	0
15	x		x			x		x		x	2	0
16	x		x			x		x		x	2	0
17	x			x		x	x			x	2	2
18	x			x		x		x		x	1	0
19	x			x		x		x		x	1	0
20	x			x		x		x		x	1	0
21		x	x			x		x		x	1	2
22		x		x		x		x		x	0	0
23		x		x		x		x		x	0	0
24		x		x		x		x		x	0	0
25		x		x		x		x		x	0	0
f	20	5	17	8	10	15	8	17	5	20		12
chyby	0		1		0		3		2			

Pro ověření, zda získaná data odpovídají škálogramu, uvádí Břicháček (1978) postup Cornellskou technikou nebo metodou pro stanovení stupně reprodukovatelnosti, kterou navrhl Goodenough a Edwards (Goodenough 1944, Edwards 1948 v Břicháček 1978).

U postupu Cornellskou technikou je třeba v každé položce stanovit bod, ve kterém se mění odpovědi z kategorie „ano“ („1“) na kategorii „ne“ („0“). Pro stanovení dělicích bodů (tabulka č.1, přechodové body označuje zvýrazněná vodorovná linka) platí následující pravidla:

- 1) data se musí uspořádat tak, aby vzniklo co nejmenší množství chyb (tj. odpovědí, které se liší od ostatních odpovědí ležících nad nebo pod dělicím bodem),
- 2) data se musí uspořádat tak, aby v žádné položce nebylo více chyb než správných odpovědí.

Počet chyb je potom zaznamenán do tabulky. Pro vyhodnocení, zda získaná data odpovídají ideální unidimenzionální škále, navrhl Guttman koeficient reprodukovatelnosti (rk), který určuje, zda zjištěný počet chyb je zanedbatelný či nikoliv, tj. zda získaná data odpovídají ideální škále:

$$rk = 1 - \frac{F}{i \cdot N}, \quad (1)$$

kde F je celkový počet chyb,

i značí počet položek,

N označuje počet testovaných osob.

Pro ideální perfektní škálu by se měla hodnota rk rovnat 1, jako přijatelné se uvádí hodnoty $rk = 0,85 - 0,90$.

Získané hodnoty se interpretují následujícími způsoby:

- 1) malý počet ($rk \geq 0,85$) chyb znamená, že škála je vyhovující;
- 2) velký počet chyb ($rk < 0,85$) může znamenat:
 - a) vyskytuje-li se většina chyb u jediné položky, pak je třeba položku vypustit a znovu vypočítat rk ;
 - b) pokud lze soubor položek rozdělit do dvou podskupin, je třeba jednu skupinu vybrat a znovu vypočítat rk ;
 - c) pokud nelze soubor položek rozdělit do dvou podskupin, ale je zřejmá společná dimenze, vytvoří se tzv. kvazi-škála, která může být použita pro analýzu latentních struktur;
 - d) soubor položek neodpovídá požadavkům škálogramu.

Čepička (2003) předpokládá, že v případě teoretických konceptů motoriky člověka by mělo platit totéž, co předpokládá Thurstone u konstrukce škál sestrojených za účelem hodnocení postojů. Škála by měla být nejen unidimenzionální, ale položky by měly sdílet také shodné lineární kontinuum, tj. položky by měly být ve škále uspořádány tak, aby jejich obtížnost stoupala ve škále postupně.

Guttmanova škála je škálou kumulativní a všechny položky škály musí diagnostikovat shodný latentní rys. Testovaná osoba by měla splnit blok položek od počátku škály až po kritický bod, který značí maximální možnou úroveň latentního rysu u testované osoby. Položky nacházející se nad kritickým bodem testovaná osoba splnit nemůže, protože nemá potřebnou úroveň latentního rysu. Poslední splněná položka tedy určuje skóre vypovídající o úrovni latentního rysu u testované osoby (Čepička, 2005).

Sestavení unidimenzionálních škál podle Guttmana naráží v praxi na značné potíže, koeficient reprodukovatelnosti byl často diskutován jako velice problematická část konstrukce škály. Nicméně základy unidimenzionálního škálování tak, jak je formuloval Guttman, jsou postačující s využitím dalších vhodných modelů (Břicháček 1978, Čepička 2003, 2005).

Čepička (2005) doporučuje ve své studii zabývající se diagnostikou motorických dovedností použít Raschův model. „Vzhledem k tomu, že prostřednictvím Raschova modelu lze vhodným způsobem odhadnout obtížnosti položek, nabízí se jeho použití v případě kvantifikace položek jako škálových hodnot. To je základním důvodem pro použití Raschova modelu,“ uvádí Čepička (2005). Použití Raschova modelu nebo dalších IRT modelů pro konstrukci Guttmanovy škály doporučuje také Massof (2011), který uvádí, že Guttmanova škála je teoretický a matematický ideál, u kterého se variabilita měření přenáší na chyby ve škále. Raschovy a IRT modely využívají tyto chyby jako prostředek pro odhad intervalů z ordinálních dat přidělených při pozorování.

4. TEORIE POLOŽKOVÝCH ODPOVĚDÍ

Teorie položkových odpovědí (*IRT – item response theory*), patřící mezi kritériální metody, zahrnuje modely určené převážně k analýze binárních dat a své uplatnění nacházela zpočátku především v psychologii. Základní koncept teorie založené na položkách prezentoval Lawley (Lawley 1943 v Baker a kol., 2004,1) a teorii rozšířil Lord (Lord 1952 v Baker a kol., 2004,1), který ukázal, že konstrukty klasické teorie testů mohou být vyjádřeny jako parametrické funkce charakteristické křivky položky. Kořeny teorie nalezneme také v Thurstonových a Guttmanových postupech škálování a v raných úvahách o nelineární faktorové analýze. V roce 1960 vyšla kniha dánského matematika George Rasche „*Probabilistic models for some intelligence and attainment tests*“ („Pravděpodobnostní modely pro některé inteligenční a výkonové testy“), ve které prezentoval tři různé modely, jeden z modelů určený položkám testu. Tento logistický model vešel ve známost pod jménem Raschův model (Baker a kol., 2004).

Do kinantropologie proniká teorie položkových odpovědí až ke konci minulého století, kdy se ve světě začaly objevovat v odborných kinantropologických časopisech příspěvky teoretického charakteru i praktického využití teorie, a to nejen ve sportovní psychologii, ale také při motorickém testování (Safrit 1992, Zhu 1996, 1998, Tanenbaum 1999). Tanenbaum (1999) se na poli sportovní psychologie přímo zabýval implementací Guttmanových myšlenek do Raschovy analýzy.

U nás se o teorii položkových odpovědí zmiňuje v 80. letech Blahuš (1981, 1985) a na to navazuje Čepička (1999), který využil teorii položkových odpovědí, konkrétně Raschův model, ke stanovení obtížnosti IOWA–BRACE testu. V roce 2002 vyšla jeho publikace „*Modely teorie položkových odpovědí v diagnostice motoriky člověka*“ a následovala praktická aplikace, konstrukce Guttmanovy perfektní škály pomocí Raschova modelu, určená k diagnostice motorických dovedností, konkrétně předpokladu pro manipulaci s míčem (Čepička, 2003). Čepičku, který působí na Západočeské univerzitě v Plzni, také následovali studenti při konstrukci škály diagnostikující předpoklady pro manipulaci s míčem ve volejbalu (Honsová 2003, Pavlíček 2003). Využití teorie položkových odpovědí v kinantropologii se zabývají i Čepičkovy příspěvky z roku 2004 a 2010.

Využití Raschovy analýzy v motorice člověka můžeme nalézt také v lékařském výzkumu, konkrétně v oblasti rehabilitace (De Souza 1999, Graffigna a kol. 2015, Sabari a kol. 2014, Van de Winckel a Feys 2006, Velozo, 2008).

4.1 Využití teorie položkových odpovědí

Teorie položkových odpovědí slouží k analýze převážně binárních dat (tj. v kinantropologii hodnocení zadaných úkolů ve smyslu 1 – 0, splnil – nesplnil, správně – chybně apod.) s cílem odhadnout úroveň latentního rysu na základě podmíněné pravděpodobnosti úspěšného splnění zadaného úkolu, tj. položky. Předpokládáme, že čím má testovaná osoba vyšší úroveň latentního rysu, tím má vyšší pravděpodobnost správného splnění testu (Čepička, 1999), respektive testovaná osoba má vyšší pravděpodobnost úspěšnosti na jednotlivé položce.

Základním stavebním kamenem teorie položkových odpovědí je tedy položka a její vlastnosti, z kterých vyplývá nezávislost na složení testu. Výsledný test je souborem samostatných položek a právě jejich nezávislost na testu má praktické důsledky (Urbánek, Šimeček, 2001):

1. Měření latentních rysů se neinterpretuje v kontextu populace, pro kterou byla provedena standardizace testu, v teorii položkových odpovědí nazývaná tzv. kalibrace položek, protože parametry položek jsou nezávislé na souboru, na kterém byly kalibrovány.
2. Z nezávislosti vlastností položek na složení celého testu vyplývá možnost variabilní délky a složení testu (i testy o několika vhodně zvolených položkách mohou mít vysokou reliabilitu).

4.1.1 *Základní předpoklady použití teorie položkových odpovědí*

Pro použití teorie položkových odpovědí musí být splněny následující požadavky (Čepička, 2002):

1. Princip unidimenzionality – princip předpokládá, že položky daného testu měří jen jednu dimenzi, tj. jednu vlastnost, mají jeden společný faktor. Unidimenzionalita se nejčastěji hodnotí prostřednictvím shody použitého modelu s daty.

2. Princip lokální nezávislosti – označovaný jako Lazarsfeldův princip lokální nezávislosti říká, že pro pevné hodnoty latentního rysu jsou pozorované proměnné pravděpodobnostně nezávislé:

- pravděpodobnost správné odpovědi na určité položce závisí pouze na úrovni latentního rysu potřebné pro správné zodpovězení položky,
- informace o úspěchu nebo neúspěchu na další položce neovlivňuje pravděpodobnost odpovědi na položce předcházející ani žádné jiné.

Princip nezávislosti lze matematicky vyjádřit následujícím vztahem:

$$P(x_i = 1, x_j = 1, x_k = 1/\theta) = P(x_i = 1/\theta) P(x_j = 1/\theta) P(x_k = 1/\theta), \quad (2)$$

kde $P(x_i = 1, x_j = 1, x_k = 1/\theta)$ značí pravděpodobnost úspěchu ve všech položkách (tj. pravděpodobnost úspěchu v položkách i, j, k podmíněnou úrovní latentního rysu θ), která je výsledkem jednotlivých pravděpodobností úspěchu.

3. Princip nelineárního typu závislosti – princip vyplývá z výše uvedených principů, tj. z faktu, že funkční vztah mezi úrovní latentního rysu a úspěšností na položce je vyjádřen pravděpodobností. Výsledkem je tedy křivka, která v žádném svém bodě nenabývá hodnot větších než 1 a menších než 0.

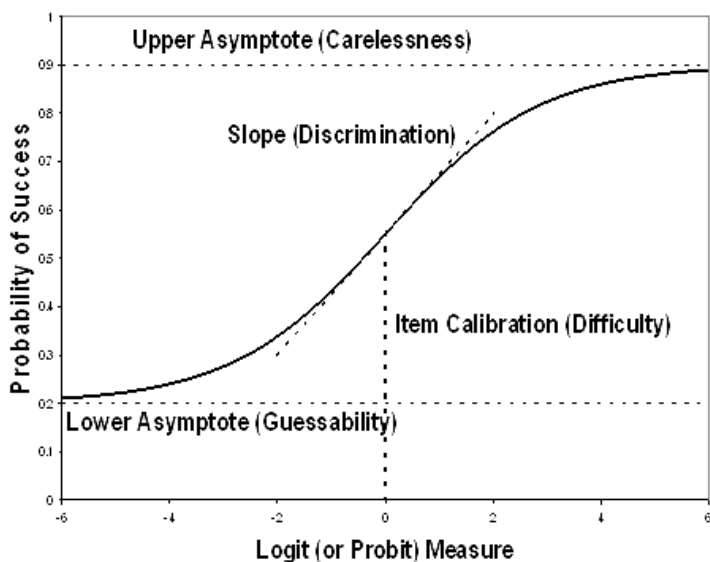
4.1.2 *Funkce položkových odpovědí*

Pomocí teorie položkových odpovědí vyjadřujeme funkční závislost – jakým způsobem úroveň latentního rysu podmiňuje správné zodpovězení položky, tj. pravděpodobnost správné odpovědi je funkcí latentního rysu. Funkci popisuje monotónně rostoucí charakteristická křivka položky (*ICC – item characteristic curve*), jejíž dolní asymptota je rovna nebo větší než 0 (v závislosti na parametru uhádnutelnosti, viz dále) a horní asymptota je rovna 1. Charakteristická křivka položky zahrnuje vztah mezi charakteristikami položky (parametry položky) a úrovní latentního rysu (*graf č.1*) a je vyjádřena v závislosti na typu položky testu (dichotomické, polytomické...) a počtu

parametrů. Odpovědi na položky jsou tedy modelovány konkrétními charakteristickými křivkami, na každou položku zvlášť, které mají v důsledku různých hodnot parametrů odlišný průběh. Při grafickém vyjádření používá teorie položkových odpovědí logitovou stupnici, tj. druh logaritmické stupnice velmi podobný z – bodové stupnici, s vlastnostmi poměrové škály (Urbánek a Šimeček 2001, Čepička 2002).

Parametry položek (Čepička, 2002):

1. Parametr diskriminativnosti – parametr diskriminativnosti formuje zakřivení funkce položkových odpovědí (*graf č.1*). Parametr charakterizuje tendenci testované osoby, která má vyšší úroveň latentního rysu, odpovídat na položku častěji správně (být častěji úspěšnější) než osoba s nižší úrovní latentního rysu.
2. Parametr obtížnosti – z *grafu č. 1* je patrné, že je parametr obtížnosti zaznamenán na stejné ose grafu jako parametr latentního rysu a posouvá křivku vůči hodnotící škále na horizontální ose. Hodnota latentního rysu a parametru obtížnosti je tedy vyjádřena na stejné stupnici. V teorii položkových odpovědí se používá zvláštní druh logaritmické stupnice – logitová stupnice. Jedná se o poměrovou stupnici nabývající hodnot $(-\infty, +\infty)$. Důležitou vlastností vzhledem k samotné teorii je fakt, že hodnota parametru je nezávislá na souboru testovaných osob (vlastnost se nazývá invariance parametrů položky).
3. Parametr uhádnutelnosti – hodnota parametru ovlivňuje polohu dolní asymptoty (*graf č.1*). Parametr uhádnutelnosti nás informuje o pravděpodobnosti, s kterou testovaná osoba s nulovou úrovní sledovaného rysu odpoví správně (bude úspěšná) na konkrétní položce. Pokud položku nelze uhádnout, pak je parametr roven 0.



Graf č. 1 Charakteristická křivka položky. Na ose x je znázorněna úroveň latentního rysu, osa y znázorňuje relativní podmíněnou pravděpodobnost správné odpovědi na položce. „*Discrimination*“ značí parametr diskriminativnosti, „*Difficulty*“ parametr obtížnosti a „*Guessability*“ parametr uhádnutelnosti. Na grafu je také znázorněna dolní a horní asymptota („*Lower/Upper Assymptote*“).

Zdroj : <http://www.rasch.org/rmt/rmt181b.htm>

Jedním z modelů teorie položkových odpovědí, který analyzuje dichotomické odpovědi na položky a předpokládá proměnlivou hodnotu pouze jednoho parametru, parametru obtížnosti, je Raschův model (parametr diskriminativnosti je pro všechny položky stejný a parametr uhádnutelnosti je roven 0).

4.2 Modely IRT

Pro klasifikaci modelů IRT jsme vybrali modely, které se již nějakým způsobem objevily nebo byly zmíněny v kinantropologickém výzkumu. Modely jsme klasifikovali podle počtu parametrů.

Pro přehlednost jsme modely zakreslili do schématu (*schéma č.1*), který zobrazuje rozdělení modelů nejen podle počtu parametrů, ale také podle způsobu odpovědi na položku. MASI modely jsme nezařadily do dichotomických ani polytomických odpovědí na položku, protože podle dostupné literatury mohou být použity v obou případech (Spray 1990, Ostini a Nering 2006).

V klasifikaci modelů podle počtu parametrů uvádíme také rozdělení modelů do skupin podle taxonomie Thissena a Steinberga (Čepička 2002, Ostini a Nering 2006).

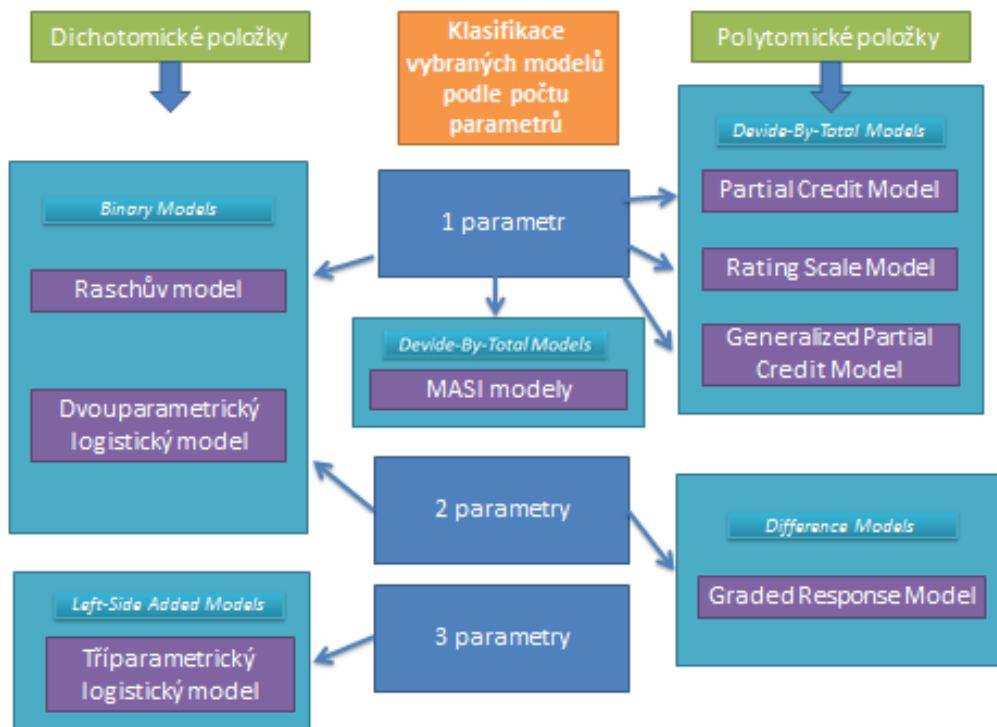


Schéma č. 1 Klasifikace vybraných modelů IRT podle počtu parametrů

Pro IRT analýzu je k dispozici specializovaný software, který je nabízený v rámci softwarových balíků „ASC“, „R“, „SAS“, „SPSS“ nebo „SSI“, jejichž součástí jsou programy pro analýzu víceparametrových IRT modelů (např. IRTPRO, BILOG–MG, PARSCALE, MULTILOG, Xcalibre nebo WINGEN). Pro analýzu jednoparametrických modelů je určen software především od autora Linacreho WINSTEPS a Facets Rach Software nebo RUMM autorů Sheridan, Andrich a Luo. Hendl (2015) doporučuje k analýze volně dostupný program WINMIRA. Massof (2011) ve své studii uvádí jako nejpoblárnější softwarové balíky pro výpočet obtížnosti položek pomocí Raschova modelu Winsteps 17 a RUMM 18 a pomocí ostatních IRT modelů programy Multilog 19 a Parscale 20.

4.2.1 *Jednoparametrické IRT modely a Raschův model*

Za nejjednodušší IRT model je považován jednoparametrický logistický model, který je popsán matematickou rovnicí (Baker a kol., 2004):

$$P_i(\theta) = \frac{1}{1 + e^{-(\theta - b_i)}}, \quad (3)$$

kde $P_i(\theta)$ je pravděpodobnost správné odpovědi na položce i při úrovni latentního rysu θ , e je Eulerovo číslo.

Každá položka je tak charakterizována pouze jedním parametrem, a to svou obtížností. Nejrozšířenějším jednoparametrickým modelem je Raschův model. Raschův model je považován za výchozí model pro další jednoparametrické „Divide–By–Total“ modely včetně MASI IRT modelů.

Raschův model

Jedním ze základních binárních modelů IRT, který analyzuje dichotomické odpovědi na položky a předpokládá proměnlivou hodnotu pouze jednoho parametru (parametru obtížnosti) je Raschův model.

Raschův model je uváděn jako zvláštní případ logistického modelu, označovaného také jako jednoduchý logistický model nebo latentní logistický model. Raschův model byl navržen pro odhad obtížnosti položek nezávisle na testovaném souboru osob a pro odhad úrovně latentního rysu u testované osoby nezávisle na položkách, na které právě odpovídala. Použití Raschova modelu má zásadní výhody. Při IRT analýze pomocí Raschova modelu jak odhadovaná úroveň latentního rysu, tak obtížnost položky jsou kvantifikovány na stejné měřicí škále (Raschův model tak poskytuje přímo vyjádřené srovnání hodnot úrovně latentního rysu a obtížnosti položky) a odhad obtížnosti položky je získáván nezávisle na distribuci latentního rysu ve výběrovém souboru. Parametr obtížnosti tak existuje nezávisle na testované osobě a poměr mezi úrovní latentního rysu a obtížností položky zůstává zachován (Čepička 2003, 2005).

Raschův model tedy modeluje vztah mezi úrovní latentního rysu a odpovědí na položku u testované osoby. Čím vyšší je úroveň latentního rysu u testované osoby než je obtížnost položky, tím je vyšší pravděpodobnost úspěšnosti testované osoby pro splnění úkolu na položce. Je-li hodnota latentního rysu nižší, úspěšnost klesá. Je-li obtížnost položky shodná s hodnotou latentního rysu, potom je pravděpodobnost rovna 50%.

Položková funkce *Raschova modelu* má tvar (Čepička, 2002):

$$P(\theta) = \frac{e^{(\theta-b)}}{1 + e^{(\theta-b)}}, \quad (4)$$

kde e je Eulerovo číslo, θ je parametr latentního rysu, b značí parametr obtížnosti.

Snadnost testové položky můžeme také vyjádřit vzorcem (Čepička, 2002):

$$e^{(\theta-b)} = \frac{P_i(\theta)}{Q_i(\theta)}, \quad (5)$$

kde e je Eulerovo číslo, $P_i(\theta)$ představuje pravděpodobnost správné odpovědi (splnění úkolu) na položce i a $Q_i(\theta)$ pravděpodobnost chybné odpovědi (nesplnění úkolu) na položce i .

Nevýhodou Raschova modelu je, že u všech položek se předpokládá stejná diskriminace (pravděpodobnost, že testované osoby, které mají vyšší úroveň latentního rysu, budou odpovídat na položku častěji správně než osoby s nižší úrovní latentního rysu, je u všech položek stejná). Raschův model také nebere v úvahu možnost uhádnutí položky (tj. pravděpodobnost, s kterou testovaná osoba s nulovou úrovní sledovaného rysu odpoví správně na konkrétní položce).

MASI IRT modely

Modely, které mohou analyzovat testy skládající se z opakujících se nebo násobících se pokusů na stejných dichotomických položkách (v kinantropologii např. střílení trestných košů v basketbale nebo test “sed-leh“ za 1 min.), tj. „*multiple attempt, single item*“, se nazývají MASI IRT modely. Uplatnění modelů v kinantropologii popsala ve své teoretické stati Spray (1990). Mezi MASI modely patří Binomial Trials model, Poisson Counts model a Inverse Binomial model.

U Binomial Trials modelu je pozorované skóre testu (x_i) výsledkem celkového počtu správných odpovědí z n možných odpovědí (pokusů) na položce. Pokud předpokládáme, že testovaná osoba může opakovat n pokusů na jedné položce a pravděpodobnost úspěšných pokusů zůstává konstantní u každého pokusu pro všechny testované osoby, pak pro celkový počet úspěšných pokusů (x_i) může být použita binomiální distribuce. Pravděpodobnost, že testovaná osoba bude mít skóre testu x_i je (Spray, 1990):

$$P(X_i = x_i/\theta_i, b) = \binom{n}{x_i} \frac{e^{x_i(\theta_i-b)}}{[1 + e^{(\theta_i-b)}]^n}, \quad (6)$$

kde e je Eulerovo číslo, θ_i je parametr latentního rysu, b značí parametr obtížnosti, n je celkový počet pokusů, který je předem dán, x_i je celkový počet úspěšných pokusů.

Poisson Counts model se používá u testů, ve kterých není předem určen počet možných pokusů na jedné položce jako u předchozího modelu nebo je dán časový interval pro opakování pokusů na položce. Pravděpodobnost, že testovaná osoba s úrovní latentního rysu θ_i bude mít skóre testu x_i , určuje vztah (Spray, 1990):

$$P(X_i = x_i/\theta_i, b) = \frac{e^{x_i(\theta_i-b)}}{x_i! e^{e^{(\theta_i-b)}}}, \quad (7)$$

kde e je Eulerovo číslo, θ_i je parametr latentního rysu, b značí parametr obtížnosti, x_i je celkový počet úspěšných pokusů.

Pokud je v testu místo předem daného počtu pokusů určen počet úspěšných pokusů, kterých testovaná osoba musí dosáhnout, pak je možné použít Inverse Binomial model. Pravděpodobnost testované osoby s úrovní latentního rysu θ_i a počtem pokusů x_i je dána vztahem (Spray, 1990):

$$P(X_i = x_i/\theta_i, b) = \binom{x_i - 1}{s - 1} \frac{e^{s(\theta_i-b)}}{[1 + e^{(\theta_i-b)}]^{x_i}}, \quad (8)$$

kde e je Eulerovo číslo, θ_i je parametr latentního rysu, b značí parametr obtížnosti, s je předem daný počet úspěšných pokusů, kterých testovaná osoba musí dosáhnout, x_i je celkový počet pokusů.

Výhody MASI IRT modelů jsou shodné s výhodami Raschova modelu (přímo vyjádřené srovnání hodnot úrovně latentního rysu a obtížnosti položky a nezávislost parametru obtížnosti na testované osobě). MASI IRT modely se využívají také k testování změn a zlepšení úrovně latentního rysu θ_i v čase (při testování osob MASI pretestem a MASI posttestem zůstává obtížnost testu b , ale mění se úroveň latentního rysu θ_i , která je odhadována pro každou testovanou osobu zvlášť).

Nevýhodou MASI IRT modelů je omezení počtu parametrů jako v případě Raschova modelu (absence parametru diskriminativnosti a parametru uhádnutelnosti).

Mezi „Divide-By-Total“ modely patří Mastersův Partial Credit model, Andrichův Rating Scale model a Generalized Partial Credit model. Jedná se o zobecněné Rasch modely určené k analýze polytomických položek. Modely mají své uplatnění především v psychometrickém měření, nicméně například Rating Scale model byl využit i k hodnocení úrovně motorické dovedností (Safrit, 1987).

4.2.2 Dvoupametrické IRT modely

Dvoupametrický model IRT rozšiřuje jednapametrický logistický model o další parametr – parametr diskriminativnosti. Dvoupametrický model je popsán matematickou rovnicí (Baker a kol., 2004):

$$P_i(\theta) = \frac{1}{1 + e^{-a_i(\theta - b_i)}}, \quad (9)$$

kde $P_i(\theta)$ je pravděpodobnost správné odpovědi na položce i při úrovni latentního rysu θ , e je Eulerovo číslo, a_i je parametr diskriminativnosti.

Mezi dvoupametrické modely patří například Graded Response Model určený ke zpracování polytomických položek. Používá se pro analýzu dat, které mohou být seřazeny do m kategorií na položce i (Samejima 1969 v Safrit, Cohen a Costa 1989). V kinantropologii použili princip Graded Response Modelu Safrit, Cohen a Costa (1989), ale pouze s jedním parametrem, parametrem obtížnosti.

Výhodou těchto modelů je přítomnost parametru diskriminativnosti, který charakterizuje tendenci testované osoby, která má vyšší úroveň latentního rysu, být na položce častěji úspěšnější než osoba s nižší úrovní latentního rysu (Čepička, 2002). Nevýhodou modelů je, že modely neberou v úvahu možnost uhádnutí položky, což může být problematické v testech u položek, kde se vybírá odpověď z několika možností.

4.2.3 *Tříparametrické IRT modely*

Tříparametrický model IRT rozšiřuje dvouparametrický logistický model vedle parametru obtížnosti a diskriminativnosti ještě o parametr uhádnutelnosti. Tříparametrický model je popsán matematickou rovnicí (Baker a kol., 2004):

$$P_i(\theta) = c_i + (1 - c_i) \frac{1}{1 + e^{-a_i(\theta - b_i)}}, \quad (10)$$

kde $P_i(\theta)$ je pravděpodobnost správné odpovědi na položce i při úrovni latentního rysu θ , e je Eulerovo číslo, a_i je parametr diskriminativnosti, c_i je parametr uhádnutelnosti.

Tříparametrový logistický model je nejvíce užívaný model v „kognitivních testech“ pro dichotomické položky (Zhu, 2006). Výhodou tříparametrických modelů je přítomnost parametru uhádnutelnosti, který nás informuje o pravděpodobnosti, se kterou testovaná osoba s nulovou úrovní sledovaného rysu bude úspěšná na konkrétní položce (Čepička, 2002). Nevýhodou modelů je značná výpočetní náročnost.

4.3 Aplikace modelů IRT v kinantropologii

IRT modely se dnes používají v mnoha souvislostech. Neprovádí se pomocí nich jen položková analýza, ale vývoj IRT otevřel nové cesty pro aplikace vyrovnávání testů (*test equating*), adaptivního testování (*adaptive testing* nebo *tailored testing*) pro výběr položek, které nejlépe odpovídají úrovni schopností nebo dovedností testované osoby, tj. hledáme pro jedince specificky nejúčinnější testy na základě výsledků prvních změřených položek, nebo například vytváření banky položek (*item bankig*) se známými charakteristikami pomocí kterých se mohou vytvářet nové testy nebo mohou být použity k adaptivnímu testování. (Hendl 2015, Thomas a kol. 2005, Spray 1989)

V kinantropologii po představení IRT (Spray, 1987) došlo na konci osmdesátých let a v devadesátých letech k několika praktickým aplikacím Raschova modelu a několika modelů vycházejících z Raschova modelu. V roce 1987 analyzoval Safrit pomocí Raschova modelu a Rating Scale modelu jednak hráčskou dovednost při bowlingu a jednak „cutoff“ skóre. Pro další analýzu dat v bowlingu, konkrétně analýzu úrovně hráčské dovednosti, Safrit, Cohen a Costa (1989) použili Graded Response model. V devadesátých letech pak odhadovali Safrit, Zhu, Costa a Zhang (1992) obtížnost několika testů „sed-leh“ pomocí Poisson Counts modelu a o čtyři roky později Zhu a Cole (1996) znovu kalibrovali diagnostický nástroj – Ulrichův motorický test. Na konci devadesátých let se teorií IRT začíná u nás zabývat Čepička (1999), který pomocí Raschova modelu kalibruje IOWA–BRACE test a na začátku nového tisíciletí pokračuje v aplikaci Raschova modelu (2003, 2005), který využívá pro konstrukci motorické škály.

Při vyhledávání studií využívající IRT v posledních deseti letech je možné objevit velké množství článků z oblasti lékařského výzkumu, rehabilitace a psychologie (De Souza 1999, Graffigna a kol. 2015, Sabari a kol. 2014, Van de Winckel a Feys 2006, Velozo, 2008), ve kterých autoři používají různé modely, nicméně aplikace IRT v kinantropologii jsou ojedinělé. V roce 2008 použil Rating Scale model Myers, Feltz a Wolfe k měření efektivity koučování trenérů mládeže na školách a v roce 2011 využívá Raschův model Zhu a kolektiv autorů, kteří vytvářejí sadu testů určených pro hodnocení studentů vzhledem k cílům národních standardů v tělesné výchově v USA. Rozsáhlá práce vyústila ve vytvoření položkové banky „PE Metrics“ určené pro učitele tělesné výchovy.

V roce 1987, když Spray publikovala první vědecký článek o IRT a její možné aplikaci při tvorbě testů v tělesné výchově, napsala: „Nevidím důvod, proč by výzkum v IRT nemohl mít ihned přínos pro učitele tělesné výchovy v oblasti testování kognitivního a emocionálního chování,“ a dále ve své práci uvedla příklady využití již existujících modelů v tělovýchovné praxi, kterým se dále věnovala i v následujícím vědeckém článku z roku 1990. Na teoretická východiska od Spray (1987, 1990) navázaly výše uvedené praktické pokusy využití IRT v tělesné výchově a sportu (např. Safrit 1987, 1989, 1992; Zhu 1996, 1998).

Kořeny IRT sahají až do roku 1940, k největšímu rozvoji došlo v šedesátých a sedmdesátých letech spolu s výpočetní technikou, nicméně v kinantropologii se teorie začala řešit na vědecké úrovni až v devadesátých letech. Spray (1987) si proto klade otázku: „Proč měla IRT až do dnešních dnů tak malý vliv na měření v tělesné výchově?“ Některé důvody uvádí v reakci na článek Wood (1987), například vysokou úroveň matematických dovedností nutnou pro studium potřebné literatury, nedostatek vhodných počítačových programů nebo neochotu měnit tradiční metody klasické teorie testů (CTT). Wood (1987) se dále vyjadřuje k možnostem v psychomotorickém testování a nutným dalším krokům ve výzkumu IRT na poli kinantropologie – porozumění současným IRT modelům, vývoji a testování nových modelů a vývoji nových počítačových programů pro odhad parametrů. I přes některé obavy oba vědecké články (Spray 1987, Wood 1987) ve výsledku vyznívají velmi odhodlaně vzhledem k aplikaci IRT v psychomotorickém testování a s nadějí celkového zlepšení testů v tělesné výchově a sportu.

O více než 20 let později Zhu a kol. (2011) uvádí: „Nové testovací teorie, koncepty a psychometrické metody (např. IRT, test equating, položková banka) rozvíjející se během několika minulých dekád mají mnoho výhod na rozdíl od předchozích teorií a metod. Navzdory jejich představení v této oblasti, nebyly teorie ale plně akceptovány v tělesné výchově...IRT byla představena v oblasti tělesné výchovy před více než 20 lety a bylo zveřejněno mnoho úspěšných aplikací (např. Looney 1997, Myers a kol. 2008, Tenenbaum 1986, Zhu a Cole 1996) ...přestože existuje důvěra v IRT a její potenciál v praxi, tělesná výchova nevyužívá výhod této ‚nové‘ testovací metody a souvisejících benefitů pro měření...“.

Počet nalezených studií týkajících se využití IRT v kinantropologii je stále vzhledem k jejím možnostem nastíněným během dvaceti let teoretického výzkumu minimální. Z výše uvedených informací vyplývá, že kinantropologie plně nevyužívá možnosti IRT, které jsou k dispozici.

5. 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.

6. DESIGN A METODOLOGIE VÝZKUMU

6.1 Vymezení teoretické oblasti výzkumu

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 ovládání 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 ovládání florbalové hole a míčku u hráčů florbalu. Zjišťování úrovně dovednosti ovládání 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 ovládání 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.

6.2 Výzkumný soubor

Výzkumný soubor byl složen z 212 hráčů vybraných florbalových družstev kategorie přípravka (6 – 8 let), elévové (9 – 10 let) a mladší žáci (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.

Přehled probandů z hlediska věkové kategorie:

- 42 hráčů kategorie přípravka (6 – 8 let),
- 141 hráčů kategorie elévové (9 – 10 let),
- 29 hráčů kategorie mladší žáci (11 – 12 let).

6.3 Výzkumné metody

K výzkumu jsme 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).

6.3.1 Expertní analýza pro posouzení obsahové validity podle Lawshe

K posouzení obsahové validity jsme použili koeficient podle Lawshe (1975). Pomocí koeficientu lze numericky vyjádřit, zda experti považují navrženou položku za podstatnou či nikoliv. Koeficient podle Lawshe se vypočítá ze vzorce:

$$CVR = \frac{n_a - \frac{N}{2}}{\frac{N}{2}}, \quad (11)$$

kde N je počet všech expertů a n_a je počet expertů, kteří označili položku za podstatnou.

Položky s indexem $CVR \geq 0$ mají podle autora koeficientu uspokojivé hodnocení, zatímco položky s $CVR < 0$ nejsou vhodné (Vágner, 2010).

6.3.2 Ztrátové fit funkce pro určení shody modelu s daty

Jednou z metod posuzující „vhodnost“ („*fit*“) u Raschových modelů, a také nejrozšířenější mezi programy používanými ke kalibraci a analýze, je založená na spojování („*concatenation*“) reziduí položek nebo osob (Smith a kol., 1995). Fit statistiky položek byly poprvé navrženy autory Wright a Panchapakesan (Wright a Panchapakesan 1969 in Smith a kol., 1995, 2) a byly založeny na hrubém skóru skupin osob, které byly zaměřeny na rozdíl mezi pozorovaným a očekávaným skóre pro skupiny osob se stejným hrubým skóre v testu. Následující výzkum ohledně fit statistik byl založen na reziduích položek a osob.

Data se ve statistice modelují podle schématu (Hendl, 2009):

data = predikce modelem (funkce) + reziduální hodnota,

kde reziduální hodnota (tzv. šum) reprezentuje diferenci mezi modelem a napozorovanými daty.

Statistiky „vhodnosti“, tzv. fit statistiky, se potom vyjadřují jako střední kvadratická rezidua nebo standardizované hodnoty rezidua.

Fit statistiky dělíme na (Čepička, 2005):

- a) outfit statistiky – jedná se o průměrnou hodnotu rozptylu standardizovaných reziduí položek i testovaných osob. Outfit statistiky akcentují neočekávané odpovědi vzdálené od vypočítaného parametru položky a jsou definovány vztahem:

$$outfit = \frac{\sum z_{ni}^2}{N}, \quad (12)$$

kde z_{ni} je standardizované reziduum pro každou osobu v každé položce.

- b) infit statistiky – jedná se o vážená standardizovaná rezidua, která jsou vážená svým vlastním rozptylem z důvodu snížení vlivu neočekávaných odpovědí vzdálených od

pozice vypočítaného parametru položky. Infit statistiky akcentují neočekávané odpovědi blízké pozici vypočítaného parametru položky a jsou definovány vztahem:

$$infit = \frac{\sum z_{ni}^2 \cdot W_{ni}}{\sum W_{ni}}, \quad (13)$$

kde W_{ni} je rozptyl testované osoby v dané položce.

6.3.3 Reliabilita škály u Raschova modelu

Pro posouzení celkové reliability testu jsme použili Kuder – Richardsonův koeficient KR – 20 (Cronbach, 1951). Koeficient Cronbachovo alfa je obecnou verzí Kuder – Richardsonova koeficientu ekvivalence. KR – 20 se používá pouze pro dichotomická data, zatímco koeficient alfa se používá na libovolný soubor položek bez ohledu na strukturu odpovědí (Cortina 1993, Cronbach 1951). Koeficient KR – 20 se vypočítá:

$$r_{tt} = \left(\frac{L}{L-1} \right) \cdot \frac{\sigma_t^2 - \sum_{i=1}^L p_i \cdot q_i}{\sigma_t^2}, \quad (14)$$

kde L je délka testu, σ_t^2 je pozorovaný rozptyl hrubého skóre testovaných osob, p_i je množství osob, které úspěšně splnily položku i , q_i je množství testovaných osob, které nesplnily položku i , součin $p_i q_i$ odhaduje rozptyl chyb.

Reliability škály u Raschova modelu můžeme dále odhadovat jednak pro testované osoby a jednak pro položky. Reliabilita odhadovaná pro testované osoby posuzuje, do jaké míry lze zachovat pořadí testovaných osob seřazených podle jejich úrovně latentního rysu v případě, že se jim předloží jiná sestava položek diagnostikující tentýž latentní rys. Naopak reliabilita položek nám říká, do jaké míry lze zachovat pořadí položek, pokud bude testován

jiný soubor osob (Čepička 2005). Obě reliability jsou ekvivalentní ke Kuder – Richardsonovu koeficientu KR – 20 (Linacre, 2017).

Reliabilita je obecně definována následujícím vztahem (Hendl, 2016):

$Reliabilita = rozptyl\ pravdivého\ skóru / (rozptyl\ pravdivého\ skóru + chybový\ rozptyl),$

kde vztah $rozptyl\ pravdivého\ skóru + chybový\ rozptyl$ značí napozorovanou hodnotu (tj. *pozorovaný rozptyl*).

Reliabilita Raschova modelu (Čepička 2005, Linacre 2017) je definována vztahem:

$Reliabilita\ RM = (pozorovaný\ rozptyl - reálný\ rozptyl\ chyb) / pozorovaný\ rozptyl,$

kde *reálný rozptyl chyb* odpovídá vztahu $modelový\ rozptyl * max.\ hodnota\ infit.$

Modelový rozptyl je odhad modelovaných rozptylů pozorovaného skóre vycházející ze vztahu

$$x_{ni} = P_{ni} \pm \sqrt{P_{ni}(1 - P_{ni})}, \quad (15)$$

kde $x_{ni} = 0; 1$ je odpověď osoby n v položce i , P_{ni} je pravděpodobnost správné odpovědi osoby n v položce i a $P_{ni}(1 - P_{ni})$ je rozptyl pozorovaného skóre v případě x_{ni} . *Max. hodnota infit*, tj. maximální hodnota středního kvadratického rezidua v infit statistice, je odhadem rušivých vlivů na data.

Reliabilita testovaných osob neboli index separace testovaných osob definujeme:

$$R_i = \frac{\sigma_{TP}^2 - \sigma_{EP}^2}{\sigma_{TP}^2}, \quad (16)$$

kde σ_{TP}^2 je celkový rozptyl skóre testovaných osob a σ_{EP}^2 znamená rozptyl chyb.

Reliabilita položek se definuje stejným vztahem s tím rozdílem, že rozptyl osob je nahrazen rozptylem položek:

$$R_i = \frac{\sigma_I^2 - \sigma_{EI}^2}{\sigma_I^2}. \quad (17)$$

U reliability raschových měření se uvádí také separační koeficient, který nabývá hodnot od 0 do nekonečna a vztah k reliabilitě osob a položek je popsán následujícím způsobem (Linacre, 2017):

$$\text{Reliabilita} = \text{separační koeficient}^2 / (1 + \text{separační koeficient}^2).$$

6.3.4 Fleissovo kappa pro zjištění shody posuzovatelů

K zjištění absolutní míry shody posuzovatelů, objektivitu diagnostického nástroje, se v současné době používá kappa koeficient (Hendl, 2015). Kappa koeficient uvedl Cohen (1960) za účelem zjišťování shody dvěma hodnotiteli s daty získanými na nominální škále. V roce 1971 byl kappa koeficient modifikován Fleissem k posouzení shody mezi více než dvěma posuzovateli (Fleiss, 1971).

Kappa koeficient je navržen pro nominální klasifikace, ke kterým se řadí i dichotomická data. Koeficient kappa se vypočítá pomocí vzorce (Hendl, 2015):

$$\kappa = \frac{p_o - p_e}{1 - p_e}, \quad (18)$$

kde p_o je odhad pravděpodobnosti, že se při klasifikaci jevu oba posuzovatelé shodnou, a p_e odhaduje náhodnou shodu posuzovatelů a vypočítá se pomocí vzorce:

$$p_e = \frac{1}{n^2} \sum_{j=1}^K n_i \cdot n_j, \quad (19)$$

kde četnosti $n_{i,j}$ označují počet objektů/jevů zařazených posuzovatelem A do i -té kategorie a posuzovatelem B do j -té kategorie. Počet shod obou posuzovatelů se získá součtem četností $n_{i,j}$.

Fleiss (1971) modifikoval vzorec k výpočtu shody mezi více než dvěma posuzovateli následujícím způsobem:

$$K = \frac{\sum_{i=1}^N \sum_{j=1}^k n_{ij}^2 - Nn \left[1 + (n-1) \sum_{j=1}^k p_j^2 \right]}{Nn(n-1) \left(1 - \sum_{j=1}^k p_j^2 \right)}, \quad (20)$$

kde N je počet subjektů (probandů), n je počet hodnocení na subjekt (počet examinátorů), k počet kategorií (položek), index i , kde $i = 1, \dots, N$, reprezentuje subjekty, a index j , kde $j = 1, \dots, k$, reprezentuje kategorie škály, n_{ij} je počet examinátorů, kteří přidělili i -tý subjekt k j -té kategorii a pravděpodobnost p_j je definována vztahem:

$$p_j = \frac{1}{Nn} \sum_{i=1}^N n_{ij}. \quad (21)$$

Výsledkem jsou hodnoty v intervalu $(-1,1)$, kde hodnota 1 reprezentuje úplnou shodu a hodnota 0 znamená žádnou shodu mezi posuzovateli, negativní hodnoty pak naznačují inverzní tvrzení (Zapf a kol., 2016).

6.3.5 Analýza hlavních komponent pro určení unidimenzionality škály

Analýza hlavních komponent („PCA – *Principal component analysis*“) byla poprvé popsána Karlem Pearsonem v roce 1901 a dále byla rozvíjena Haroldem Hotellingem v třicátých letech minulého století. Analýza je založena na zkoumání hodnot nových proměnných (hlavních komponent), které vznikají jako lineární kombinace původních proměnných a umožňují nám tak lépe porozumět zkoumaným datům (Hendl, 2015). Analýza hlavních komponent je vícerozměrná metoda, která přeměňuje korelované proměnné na nekorelované proměnné pomocí zmenšení rozptylu (Wright, 1996).

Při analýze vycházíme z předpokladu, že pouze několik proměnných, hlavních komponent, má nezanedbatelný rozptyl a ostatní proměnné můžeme tak při analýze zanedbat. Dosáhne se tak úspornějšího popisu chování původních proměnných pomocí menšího počtu nových proměnných, hlavních komponent. Obvykle má větší rozptyl pouze několik prvních hlavních komponent, které pak slouží k rekonstrukci všech ostatních proměnných. Celkový rozptyl lze zachytit pouze jednou hlavní komponentou, pokud všechny proměnné korelují s korelací blízkou jedné v absolutní hodnotě. (Hendl, 2015)

U Raschova modelu použijeme analýzu hlavních komponent reziduí, jejímž cílem je vyvrácení hypotézy, že rezidua jsou náhodným šumem. Snažíme se o to pomocí nalezení komponenty, která vysvětluje největší možný rozptyl v reziduích u dat, která se neshodují s Raschovým měřením. První komponenta (Raschova dimenze) má největší rozptyl a je u analýzy hlavních komponent reziduí odstraněna a hledáme tedy až druhou dimenzi, komponentu, která vysvětluje rozptyl u prvního kontrastu. Komponenta ukazuje kontrast mezi opačnými vektory, na rozdíl od faktorové analýzy, která zjišťuje faktorové zátěže. Pokud se rozptyl nachází nad stanovenou hranicí šumu, existuje druhá dimenze a pokračujeme ve hledání třetí dimenze atd. (Linacre, 2017)

Výše uvedené vztahy se popisují jednak procentuálně, kolik procent rozptylu je vysvětleno pomocí dané komponenty, a také pomocí tzv. *eigenvalue*, vlastního čísla, odpovídajícího tzv. *eigenvektoru*, které odpovídá hodnotě velikosti rozptylu v datech okolo dané komponenty.

Matematicky můžeme analýzu hlavních komponent reziduí vyjádřit následujícími vztahy (Linacre 2017, Write a Masters 1982):

$$\text{Pozorovaný rozptyl v datech: } OV = \sum_{n=1}^N \sum_{i=1}^L (X_{ni} - A)^2, \quad (22)$$

$$\text{Rozptyl vysvětlený měřením: } RV = \sum_{n=1}^N \sum_{i=1}^L (E_{ni} - A)^2, \quad (23)$$

$$\text{Předpokládaný (modelový) rozptyl: } MV = \sum_{n=1}^N \sum_{i=1}^L W_{ni}, \quad (24)$$

$$\text{Nevysvětlitelný rozptyl: } UV = \sum_{n=1}^N \sum_{i=1}^L (X_{ni} - E_{ni})^2, \quad (25)$$

$$\text{Celkový rozptyl v datech: } TV = RV + MV, \quad (26)$$

$$\text{Poměr mezi rozptyly: } RV / TV, \quad (27)$$

$$\text{Pokud data odpovídají modelu, potom platí vztah: } OV = TV, \quad (28)$$

Protože data nikdy ideálně neodpovídají modelu, rozptyl je popsán vztahem:

$$AV = OV + UV, \quad (29)$$

kde N je počet probandů odpovídajících na L dichotomických položkách, výsledek odpovědi osoby n na položku i vyjadřuje X_{ni} , jehož očekávanou hodnotou je E_{ni} a rozptyl predikovaný modelem je W_{ni} . A je definováno vztahem:

$$A = \sum_{n=1}^N \sum_{i=1}^L X_{ni} / NL. \quad (30)$$

6.4 Postup při konstrukci motorické škály

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

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*.

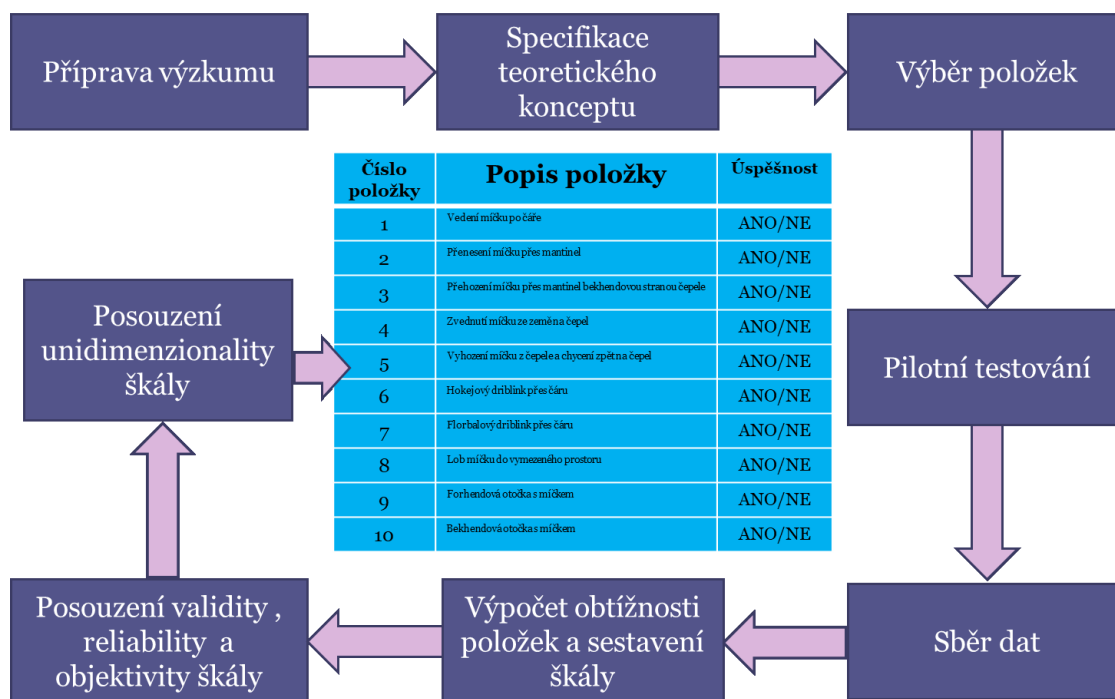


Schéma č. 2 Postup při standardizaci škály pro posouzení úrovně florbalové dovednosti

6.4.1 Příprava výzkumu

V této části výzkumu je důležité si představit celkový design výzkumu, proces, finanční zajištění, materiální a časovou náročnost, metody vhodné pro statistickou analýzu apod. (Štochl a Musálek, 2009). V této části výzkumu jsme se proto zabývali:

- výběrem expertů pro hodnocení obsahové validity škály,
- výběrem testovaných osob,
- výběrem místa pro testování,
- sestavením harmonogramu výzkumu,
- výběrem examinátorů pro jednotlivé fáze testování,
- výběrem examinátorů pro hodnocení objektivit škály,
- výběrem statistických metod,
- výběrem statistického programu.

6.4.2 Specifikace teoretického konceptu

Štochl a Musálek (2009) upozorňují na problematiku specifikace teoretického konceptu, která se může zdát jednoduchou, ale mnoho studií, jak autoři uvádějí, bylo právě vzhledem ke špatné specifikaci teoretického konceptu neúspěšných. Pro úspěšné dokončení celého procesu standardizace považují tento krok za nejdůležitější a doporučují vědeckým pracovníkům spolupráci s dalšími odborníky – metodology, statistiky, specialisty na danou problematiku apod.

V našem výzkumu jsme tento krok po prostudování dostupné literatury konzultovali se dvěma florbalovými specialisty a metodologem. Zjišťovat budeme úroveň latentní proměnné: „florbalová dovednost, která determinuje možnosti hráče při manipulaci s míčkem“.

6.4.3 Výběr položek

Při výběru položek se doporučuje volit větší počet položek (alespoň 12), protože „čím více položek je k dispozici, tím spíše se podaří z nich vybrat část, která splňuje požadavky škálogramu...“ (Břicháček, 1978). Štochl a Musálek (2009) navrhují sepsání kompletního seznamu relevantních položek, které měří dané koncepty ve vymezené oblasti, a jejich podstoupení expertům a následné vytvoření první verze testu z položek, které experti hodnotili jako vhodné nebo nezbytné.

Pro naše účely jsme navrhli 30 položek (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 čárami.
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.

Tři položky jsme vybrali ze standardizovaných testů pro dospělé hráče odvozených z pozemního hokeje od Tůmy (2004), šest položek jsme zvolili z baterie testů švédské florbalové unie (Karlberg, 2013) a ostatní položky jsme navrhli na základě vlastních florbalových zkušeností a z cvičení používaných při florbalovém tréninku (Zlatník 2004, Skružný 2005, Kysel 2010, Martínková 2009).

Při tvorbě položek jsme vycházeli ze šesti skupin navržených Měkotou a Blahušem (1983) pro testování dílčích dovedností v míčových hrách. Snažili jsme se o rozložení položek mezi uvedené skupiny, které jsme převedli na činnosti s florbalovou holí a míčkem:

- a) stěna a míč – testy s použitím mantinelu – položka č. 19, 20, 21;
- b) hod (kop) míčem na cíl – střelba na cíl – položka č. 22, 25, 28, 29;
- c) lokomoce s míčem – testy s vedením míčku a slalomové dráhy – položka č. 1, 2, 9, 10, 11, 12, 13;
- d) hod (kop) míčem na vzdálenost – lob na vzdálenost – položka č. 30;
- e) žonglování míčem – testy zaměřené na manipulaci s florbalovou holí a míčkem – položka č. 3, 4, 5, 6, 7, 8, 14, 15, 16, 17, 18;
- f) řetězec pohybových činností s míčem – testy zahrnující několik navazujících florbalových činností za sebou – položka č. 23, 24, 26, 27.

Navržené položky jsme předložili 5 florbalovým expertům (metodici České florbalové unie, trenéři s nejvyšší trenérskou licenci a trenéři mládeže), kteří položky hodnotili z hlediska jejich obsahu vzhledem k testování florbalové dovednosti ovládní florbalovou holí a míčku u hráčů florbalu mladšího věku. Experti položky rozdělili na podstatné a nepodstatné a pomocí expertní analýzy, koeficientu podle Lawshe (1975), jsme položky seřadili. Ze skupiny položek s nejvyšším koeficientem jsme vytvořili první verzi testů a přistoupili k pilotnímu testování.

6.4.4 Ověření položek při pilotním testování

Smyslem této fáze je podle Štochla a Musálka (2009) ověřit, zda je test adekvátně obtížný pro cílovou skupinu. Výsledkem je zpětná vazba od probandů, která by měla následně vést k případnému vypuštění nevhodných položek. Velikost vzorku nemusí být příliš vysoká, obvykle stačí 20 probandů.

Pilotního testování se zúčastnilo 24 hráčů ve věku 6 – 12 let v červenci 2015 během letního florbalového soustředění.

6.4.5 Hlavní sběr dat

Cílem hlavního sběru dat je získání dat na dostatečně velkém vzorku probandů za účelem následného výpočtu obtížnosti položek a fit statistik a dále pro posouzení základních vlastností testů, jako jsou validita a reliabilita. Obecně lze doporučit nejméně deset probandů na každou jednotlivou položku testové baterie. Před samotným sběrem dat je nutné stanovit podmínky, za kterých budou data sbírána. Pro účely standardizace posuzovací škály je nutné zajistit více examinátorů, alespoň tři, pro nezávislé posuzování probandů (Štochl a Musálek, 2009).

Hlavní sběr dat se uskutečnil ve 2 fázích, v červenci a srpnu roku 2016 a následně roku 2017. Celkem bylo v obou fázích otestováno 212 hráčů mladšího školního věku.

Harmonogram testování – 1. fáze (červenec - srpen 2016):

- 42 hráčů z Čelákovic ve věku 6-12 let (*Zákupy*)
- 51 hráčů z Čelákovic, Milovic a Brandýsa n/L ve věku 6-12 let (*Čelákovice*)

Harmonogram testování – 2. fáze (červenec - srpen 2017):

- 44 hráčů z Čelákovic ve věku 6-12 let (*Zákupy*)
- 75 hráčů z Čelákovic, Milovic a Brandýsa n/L ve věku 6-12 let (*Čelákovice*)

Sběr dat probíhal ve dvou fázích, celkem 4 testovací dny, při dodržení následujících podmínek:

- všechna testování probíhala během července a srpna na florbalovém soustředění nebo kempu vždy pátý den dopoledne,
- před testy se hráči rozcvičili a provedli i rozcvičení s florbalovou holí a míčkem,
- testování probíhalo ve dvou tělocvičnách, v Zákupcích a v Čelákovících, povrch obou tělocvičen tvoří parkety,
- všichni hráči absolvovali testy v sálové sportovní obuvi a s vlastní florbalovou holí,
- hráči si jednotlivé testy před samotným testováním nezkoušeli, pro každý test je určen počet možných pokusů v případě počátečního neúspěchu.

Skupina 29 hráčů z druhé fáze byla otestována 4 examinátory pro hodnocení škály z hlediska objektivit. Examinátory byli florbaloví trenéři s délkou trenérské praxe 15 let, 9 let, 6 let a 3 roky. Různá délka trenérské praxe trenérů byla zvolena z důvodu posouzení objektivit škály při jejím budoucím použití trenéry s různou úrovní trenérských zkušeností.

6.4.6 Výpočet obtížnosti položek a sestavení škály

K sestavení škály jsme použili postup konstrukce perfektní škály pro diagnostiku motorických dovedností podle Čepičky (2003, 2005). Sestavení škály probíhá ve třech krocích:

1. vyřazení nevyhovujících položek

Vhodnost položek se posuzuje na základě jejich unidimenzionality. K posouzení použijeme *infit* a *outfit* statistiky pro jednotlivé položky, které by měly mít hodnoty v intervalu (0,7;1,3), respektive (0,5;1,5), a měly by posoudit vhodnost jednotlivých položek vzhledem k modelu. Hodnotit bychom měli také parametry obtížnosti, jejichž hodnoty by se měly pohybovat v intervalu od -3 logitů do +3 logitů. Nevyhovující položky výše uvedeným požadavkům by měly být vyřazeny.

2. kvantifikace obtížnosti jednotlivých položek

Informace o parametru obtížnosti jednotlivých položek získáme z hrubých skóreů Raschovou analýzou. Položky jsme seřadili podle velikosti od nejjednodušší po nejobtížnější a vytvořili jsme grafickou podobu škály k přehlednému zhodnocení rozložení škálových hodnot. Kvantifikace latentního rysu probíhá na shodné škále, úroveň latentního rysu odpovídá parametru obtížnosti u splněné položky.

3. zhodnocení rozložení škálových hodnot

Rozložení škálových hodnot je potřeba zkoumat z hlediska distance, která by neměla být příliš velká z důvodu ztráty diskriminativní schopnosti. Položky by měly také pokrývat kontinuum diagnostikovaného rysu v dostatečném rozpětí. Hodnota obtížnosti položky u normálně rozloženého latentního rysu by měla být v intervalu od -3 logitů do 3 logitů s pravděpodobností 95 %. Pokud rozložení škálových hodnot neodpovídá výše uvedeným požadavkům, škála by měla být doplněna o chybějící položky.

6.4.7 Posouzení validity, reliability, objektivity a unidimenzionality 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. probandů 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 diskriminalitou), pak zároveň musíme akceptovat položky s nízkou hodnotou (tj. vysokou diskriminalitou). 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 měřicího přístroj.
1,5 – 2,0	Hodnota parametru nevhodná pro konstrukci měřicího nástroje, ale nedegraduje měřicí nástroj.
0,5 – 1,5	Vhodná hodnota pro měření.
< 0,5	Hodnota parametru méně vhodná pro konstrukci měřicího nástroje, ale nedegraduje měřicí 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.

6.4.8 Zhodnocení kvality škály

Závěrečným krokem u konstrukce škály je zhodnocení její kvality na základě získaných dat. Pokud kvalita škály, respektive výsledky standardizace, neodpovídají výše uvedeným vědeckým předpokladům, je třeba se vrátit ke kroku „*Výpočet obtížnosti položek a sestavení škály*“ a pokusit se vybrat jiné položky. Je možné, že některé položky budou mít stejnou, nebo obdobnou, obtížnost a vhodné parametry fit funkcí, ale nebyly vybrány do výsledné škály. Důvodem bylo zvolení pouze jedné položky ze dvou stejně obtížných alternativ. Potom můžeme tyto vyřazené položky zařadit do nové škály, nahradit jimi některé z původních položek, a znovu provést vědeckou standardizaci viz *schéma č. 3*.

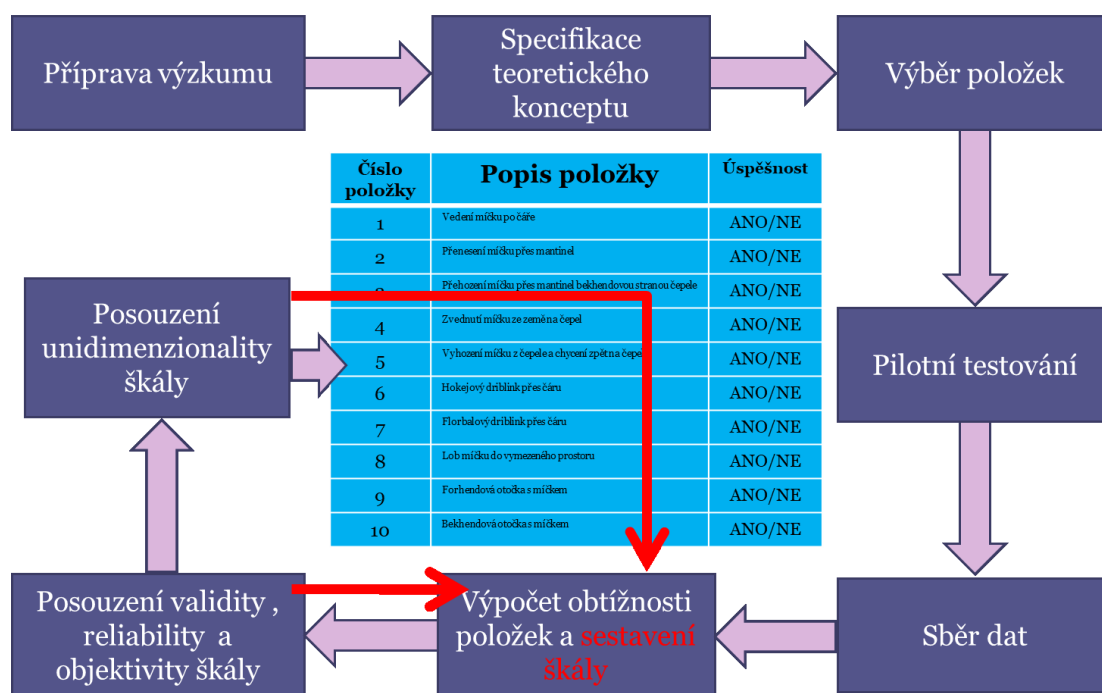


Schéma č. 3 Postup při standardizaci škály pro posouzení úrovně florbalové dovednosti, pokud výsledky standardizace nesplňují vědecká doporučení.

7. VÝSLEDKY

7.1 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.

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.

Na základě diskuse s experty jsme se rozhodli nezařadit výše uvedené položky s vyššími koeficienty do sestavy položek určené k testování z následujících důvodů:

- položka č. 5 – položka je obdobná z hlediska techniky provedení jako položka č. 6, která má vyšší koeficient;

- položka č. 12 – položku jsme vyloučili z důvodu jejího obsahu, položku z velké míry ovlivňuje úroveň pohybových schopností, úroveň pohybové dovednosti ovlivňuje úspěšnost na položce v menší míře;

- položka č. 17 – položka je obdobná z hlediska techniky provedení jako položka č. 16, která má vyšší koeficient;

- položka č. 18 – položka je obdobná z hlediska techniky provedení jako položka č. 16, která má vyšší koeficient;

- položka č. 24 – položka je obdobná z hlediska techniky provedení jako položka č. 23, která má vyšší koeficient, a položka je jednodušší z hlediska organizace testování;

- položka č. 27 – položka je obdobná z hlediska techniky provedení jako položka č. 26, která má vyšší koeficient, a položka je jednodušší z hlediska organizace testování;

- položka č. 28 – položka je obdobná z hlediska techniky provedení a zaměření (střelby na cíl) jako položka č. 25, která má vyšší koeficient.

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)

7.2 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.

Obsahem položek je hodnocení florbalového a hokejového driblinku, který provádějí probandi přes širokou čáru. Bylo velice obtížné rozdělit provedení probandů do dvou kategorií – splnil a nesplnil. Někteří probandi úlohu splnili, míček byli schopni ovládat a posouvat přes širokou čáru, ale technické provedení činnosti bylo v zásadním rozporu se správnou florbalovou technikou uváděnou u této činnosti (Kysel 2010, Martínková 2009, Skružný 2005). Jednalo se především o výrazné oddálení florbalové hole od míčku nebo o údery florbalovou holí do podlahy.

Na základě pilotního testování jsme položku č. 3 a č. 4 vyloučili ze sestavy položek a pro hlavní testování jsme využili zbylých 16 položek, které jsme přechíslovali a zjednodušili jejich název, respektive popis (*tabulka č. 6*).

Po pilotním testování jsme také ustanovili časový limit pro úspěšné splnění testu u položek obsahujících florbalový slalom, tak aby odpovídal možnostem hráčů mladšího školního věku. Časový limit pro splnění položky u florbalového slalomu „slalom s míčkem č. 1“ byl ustanoven na 10 s, u položky „slalom s míčkem č. 2“ na 13 s a u položky „slalom s míčkem č. 3“ na 12 s.

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	přihrávka z místa na cíl
Položka č.11	přihrávka 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

7.3 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.

7.3.1 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.

7.3.2 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.

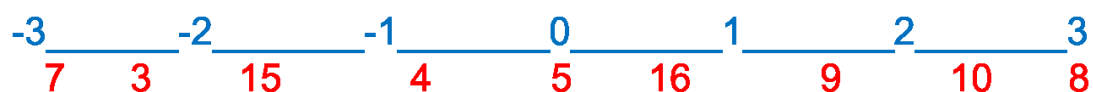
7.3.3 3. fáze – zhodnocení rozložení škálových hodnot

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 znázorněná 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

7.4 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.*

7.5 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

7.6 Standardizace škály – objektivita

Objektivitu škály hodnotili čtyři examinační 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 examinačních 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

8. 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 následujícím způsobem, podle doporučení McCreary a kol. (2014) a Linacreho (2017):

– **položka č. 1** „Forhendová otočka s míčkem“ – Položka nebyla do výsledné škály zařazena i přes vhodné hodnoty ztrátových fit funkcí (infit MNSQ = 0,93; outfit MNSQ = 1,05; infit ZSTD = -0,1; outfit ZSTD = 0,3). Důvodem byl nevyhovující parametr obtížnosti (-5,61), položka byla příliš jednoduchá.

– **položka č. 2** „Bekhendová otočka s míčkem“ – Položka nebyla do výsledné škály zařazena z důvodu vysoké hodnoty outfit MNSQ (2,49), která zbavuje hodnoty měřícího nástroje, vysoké hodnoty outfit ZSTD (2,2), která vypovídá o nedostatku předvídatelnosti a i nevyhovujícího parametru obtížnosti (-3,20), odkazujícího na přílišnou jednoduchost položky.

– **položka č. 3** „Osmička“ – Položka byla zařazena do výsledné škály, hodnota parametru obtížnosti -2,38 se nachází v intervalu (-2, -3) a hodnoty všech fit statistik se nacházejí v požadovaných intervalech (infit MNSQ = 1,01; outfit MNSQ = 1,13; infit ZSTD = 0,1; outfit ZSTD = 0,4).

– **položka č. 4** „Zvednutí míčku na čepel“ – Položka byla zařazena do výsledné škály, hodnota parametru obtížnosti -0,93 se nachází v intervalu (-1,0) a hodnoty infit a outfit MNSQ se nacházejí v požadovaném intervalu (infit MNSQ = 0,85; outfit MNSQ = 0,57). Hodnoty infit a outfit ZSTD statistik dosahují shodně hodnoty -1,6, která indikuje vyšší předvídatelnost dat, ale není hrozbou vůči validitě měřícího nástroje.

– **položka č. 5** „Vyhození míčku z čepele a chycení zpět“ – Položka byla zařazena do výsledné škály. Hodnota parametru obtížnosti -0,15 se nachází ideálně v blízkosti středu škály. Hodnota infit MNSQ 0,73 se nachází v požadovaném intervalu, ale hodnota outfit MNSQ 0,49 mírně přesahuje interval. Nicméně pro potřeby dalšího zpracování škály jsme se rozhodli tuto hodnotu akceptovat. Příliš nízké hodnoty infit a outfit ZSTD statistik (infit

ZSTD = -3,0; outfit ZSTD = -2,1) indikují vyšší předvídatelnost dat, ale nejsou jako u položky č. 4 hrozbou vůči validitě měřicího nástroje.

– **položka č. 6** „Přenesení míčku přes mantinel“ – Položka nebyla do výsledné škály zařazena i přes téměř ideální hodnoty ztrátových fit funkcí (infit MNSQ = 1,0; outfit MNSQ = 0,87; infit ZSTD = 0,0; outfit ZSTD = -0,1). Důvodem byl nevyhovující parametr obtížnosti (-3,26), položka byla příliš jednoduchá.

– **položka č. 7** „Přehození míčku přes mantinel bekhendovou stranou čepele“ – Položka byla zařazena do výsledné škály, hodnota parametru obtížnosti -2,85 se nachází v intervalu (-2, -3) a také hodnoty infit a outfit MNSQ se nacházejí v požadovaném intervalu (infit MNSQ = 0,92; outfit MNSQ = 0,58). Hodnoty infit a outfit ZSTD statistik dosahují hodnoty -0,6, respektive -0,8, a nejsou tak důvodem k vyřazení položky ze škály.

– **položka č. 8** „Vedení míčku po čáře“ – Položka byla zařazena do výsledné škály, hodnota parametru obtížnosti je 2,90 a blíží se tak krajní hodnotě škály (+3 logity). Hodnoty infit a outfit MNSQ a outfit ZSTD se nacházejí v požadovaném intervalu (infit MNSQ = 1,48; outfit MNSQ = 1,14; outfit ZSTD = 0,4)., hodnota infit ZSTD dosahuje vysoké hodnoty 3,1, která nás upozorňuje na nedostatek předvídatelnosti u položky, ale nesnižuje hodnotu škály.

– **položka č. 9** „Lob míčku do vymezeného prostoru“ – Položka byla zařazena do výsledné škály, hodnota parametru obtížnosti 1,61 se nachází v intervalu (1,2). Hodnoty infit a outfit MNSQ a ZSTD dosahují velice dobrých hodnot (infit MNSQ = 1,01; outfit MNSQ = 0,86; infit ZSTD = 0,2; outfit ZSTD = -0,2).

– **položka č. 10** „Přihrávka z místa na cíl“ – Položka byla zařazena do výsledné škály, hodnota parametru obtížnosti 2,30 se nachází v intervalu (2,3). Hodnoty infit a outfit MNSQ a outfit ZSTD se nacházejí v požadovaném intervalu (infit MNSQ = 1,18; outfit MNSQ = 1,04; outfit ZSTD = 0,2), hodnota infit ZSTD dosahuje hodnoty 1,4, poukazující na mírný nedostatek předvídatelnosti u dat.

– **položka č. 11** „Přihrávka z pohybu na cíl“ – Položka nebyla do výsledné škály zařazena i přes relativně vhodné hodnoty ztrátových fit funkcí (infit MNSQ = 0,95; outfit MNSQ = 1,19; infit ZSTD = -0,2; outfit ZSTD = 0,5). Důvodem byl nevyhovující parametr obtížnosti (3,72), položka byla příliš obtížná.

– **položka č. 12** „Střelba z místa na cíl“ – Položka nebyla do výsledné škály zařazena z důvodu nízké hodnoty outfit MNSQ (0,31) a přílišné obtížnosti položky (3,41). Nízké

hodnoty outfit MNSQ nedegradují měřicí nástroj, ale mohou vytvářet zavádějící hodnoty reliability a separece položky.

– **položka č. 13** „Střelba z pohybu na cíl“ – Položka nebyla do výsledné škály zařazena z důvodu nízké hodnoty outfit MNSQ (0,25), stejně jako u položky č. 12, a také z důvodu nevyhovujícího parametru obtížnosti (4,39), který poukazuje na přílišnou obtížnost položky.

– **položka č. 14** „Slalom s míčkem č. 1“ – Položka byla vyřazena i přes vhodný parametr obtížnosti (1,52) a relativně přijatelné hodnoty ztrátových fit funkcí (infit MNSQ = 0,78; outfit MNSQ = 0,49; infit ZSTD = -2,1; outfit ZSTD = -1,4). Důvodem byla obdobná obtížnost položky č. 9 („Lob míčku do vymezeného prostoru“), která byla do výsledné škály vybrána nejen z důvodu velice dobrých hodnot fit funkcí, ale také z hlediska obsahu položky. Obsahem položky č. 14 je slalomová dráha, která se ve výsledné škále objevuje v různých modifikacích u dvou jiných položek (č. 15 a č. 16). Z tohoto důvodu jsme upřednostnili položku č. 9, jejímž obsahem je „lob“ míčku.

– **položka č. 15** „Slalom s míčkem č. 2“ – Položka byla zařazena do výsledné škály, hodnota parametru obtížnosti -1,87 se nachází v intervalu (-1, -2). Hodnoty infit a outfit MNSQ a outfit ZSTD se nacházejí v požadovaném intervalu (infit MNSQ = 1,25; outfit MNSQ = 1,26; outfit ZSTD = 0,7), hodnota infit ZSTD dosahuje hodnoty 2,1, poukazující na nedostatek předvídatelnosti u dat, ale nezkresluje výslednou škálu.

– **položka č. 16** „Slalom s míčkem č. 3“ – Položka byla zařazena do výsledné škály, hodnota parametru obtížnosti 0,41 se nachází v intervalu (0,1). Hodnoty infit a outfit MNSQ a outfit ZSTD dosahují požadovaných hodnot (infit MNSQ = 1,14; outfit MNSQ = 0,97; outfit ZSTD = 0,0). Zvýšená je jen hodnota infit ZSTD (1,14) jako u předchozí položky.

Následně byla výsledná devítipoložková škála 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 $\kappa = 1$.

9. 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 (podrobný popis viz *příloha č. 12*):

- 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.

10. 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. JANSA, P. *Využití škálovacích technik při posuzování pohybových dovedností v tělesné výchově*. Praha, 1975. Kandidátská disertační práce na FTVS UK.
47. JANSA, P. Mnohorozměrové škálování v tělesné výchově a sportu. *Teorie a praxe tělesné výchovy*, 1989, roč. 37, č. 2.
48. JANSA, 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.

11. PŘÍLOHY

Příloha č. 1	Žádost o vyjádření Etické komise UK FTVS.....	118
Příloha č. 2	Informovaný souhlas zákonného zástupce probanda.....	119
Příloha č. 3	Matice skóru – výsledky 212 probandů na 16 položkách.....	120
Příloha č. 4	Raschova analýza 212 probandů na 16 položkách.....	125
Příloha č. 5	Raschova analýza 208 probandů na 16 položkách.....	126
Příloha č. 6	Raschova analýza 208 probandů na 9 položkách.....	127
Příloha č. 7	Grafické znázornění škály.....	128
Příloha č. 8	Analýza hlavních komponent standardizovaných reziduí.....	129
Příloha č. 9	Výstup pro výpočet reliability škály.....	130
Příloha č. 10	Data pro výpočet objektivit škály – shoda 4 posuzovatelů u 29 probandů.....	131
Příloha č. 11	Přehled 30 výchozích položek.....	134
Příloha č. 12	Pokyny k testování a popis testů.....	135