Charles University
Faculty of Physical Education and Sport

Assessment of Individual Game Skills of Ice Hockey Players on the Example of Puck Control

Abstract of dissertation

Supervisor:
doc. PaedDr. Tomáš Perič, Ph.D.

Author:
Mgr. Lukáš Chmelíř

Prague, 2025

Abstrakt

Název: Diagnostika herních činností jednotlivce u hráčů ledního hokeje na příkladu uvolňování hráče s kotoučem

Cíle: Prvním cílem práce bude vytvoření nástroje k diagnostice herních činností jednotlivce v ledním hokeji.
Druhým cílem poté bude pomocí tohoto nástroje pro hráče ledního hokeje definovat standardy v oblasti uvolňování hráče s kotoučem.

Metody: V naší práci jsme použili teorii položkových odpovědí, konkrétně Raschovu analýzu a Mokkenovu analýzu, patřící mezi neparametrické modely teorie položkových odpovědí. Pomocí Raschovy analýzy jsme vybírali položky z položkové banky do diagnostického nástroje. Reliabilita byla vypočtena pomocí Cronbachovy alfy. Pomocí Mokkenovy analýzy jsme potvrzovali výsledky Raschovy analýzy a hodnoty reliability.

Výsledky: Byl vytvořen diagnostický nástroj zahrnující 17 položek. Diagnostický nástroj je sestaven do podoby Guttmanovy škály na základě obtížnosti položek. Pomocí vytvořeného diagnostického nástroje byly definovány standardy v oblasti uvolňování hráče s kotoučem v ledním hokeji pro hráče ve věku 6–15 let. Vytvořená metodika pro tvorbu diagnostického nástroje je vhodná k použití pro tvorbu diagnostických nástrojů pro posuzování dalších dovedností v ledním hokeji.

Klíčová slova: Raschův model, Guttmanova škála, Mokkenova analýza, položky.

Abstract

Title: Assessment of individual game skills of ice hockey players on the example of puck control

Aims: The first aim of the thesis was to create a tool to assess an individual player's skills in ice hockey.
The second aim was to use the developed assessment tool to define standards for ice hockey players in terms of puck control.

Methods: This thesis employed item response theory, specifically, Rasch analysis and Mokken analysis, which belong among non-parametric models of item response theory. Rasch analysis was used to select items from the item bank for the assessment tool. The reliability was calculated using Cronbach's alpha. Mokken analysis was then used to confirm the results of Rasch analysis and reliability values.

Results: An assessment tool consisting of 17 items was developed. The diagnostic tool is constructed in the form of a Guttman scale based on the difficulty of the items. The diagnostic tool was used to define standards in terms of puck control in ice hockey for players aged 6–15 years. The methodology created for the development of the assessment tool is suitable for use in the development of diagnostic tools for the assessment of other ice hockey skills.

Keywords: Rasch model, Guttman scale, Mokken analysis, Items.

TABLE OF CONTENTS

1. INTRODUCTION

Assessment methods are very common in ice hockey. Players are tested regularly during the off-season and pre-season, in some cases, tests are also employed during the season (Bournival et al., 2023; Douglas et al., 2022; Martini et al., 2018; Pavliš et al., 2003; Perez et al., 2022; Vigh-Larsen et al., 2019). The main purpose of assessment methods is to evaluate players' readiness for the ice hockey season by checking their training. Tests of psychological aspects of training are used less frequently. The emphasis is therefore on assessing motor skills, but motor skills are only a prerequisite for performance in ice hockey. It can be said that ice hockey consists solely of skills, and there are a number of them. However, evaluating skills is less common. Moreover, players' skills change during ontogeny as their body proportions change. The degree of skill mastery also fluctuates due to the dynamics of skill learning and acquisition.

The monitoring of players' sporting performance is mainly conducted through training matches. Methods of direct and indirect observation, along with resulting expert analysis, are utilised (Tarter et al., 2017). However, these analyses are often biased (Nadeau et al., 2008; Røsten et al., 2024) by the personal experience and knowledge of the expert (Bae et al., 2024; Charbonnet & Conzelmann, 2024; Matthews & Erickson, 2024; Raudasoja et al., 2024; Urquhart et al., 2020; Wiseman et al., 2014). Especially during the school and adolescent period, coaches have varying demands on players when assessing their performance. In fact, standardised scales for assessing ice hockey skills are lacking in the literature.

Feedback is essential for effective player development, not only in ice hockey. Bukač (2014) adds that Coaching and technique adjustment are matters requiring expertise. The basis of a skill lies in its efficiency, which can be enhanced through training and mastery, and for older players, it can be regulated by the quality of training. The speed and tempo of the game are contingent on how well the player controls the puck.

The aim of this dissertation is to develop a scientifically based assessment tool for evaluating ice hockey players between the ages of 6 and 15 years. Given the large number of skills required in ice hockey, puck control was chosen as the subject of the assessment, as it is one of the fundamental skills (Český hokej, 2018; Hockey Canada, 2018; USA Hockey, 2024). This tool will be employed to define standards regarding player puck control in ice hockey for players aged 6–15 in the Czech Republic.

The outcomes of this research will contribute to a deeper understanding of motor learning and its principles. Kostka (1963) asserts that the methodology of training should continually evolve through scientific knowledge.

2. LITERATURE REVIEW

2.1. Ontogenetic Basis for Assessment Tool Development

In ice hockey, an age-appropriate training approach is recommended, starting in childhood, taking into account anatomical, physiological, educational-psychological, and individual differences. The foundation is appropriate training in childhood, followed by peak training and performance in adulthood. It is essential to respect sensitive periods in the development of motor skills. In practice, the first stages of the training process in ice hockey begin at the end of preschool and the beginning of school age. Perič & Dovalil (2010) lists the basic principles of long-term training as the principles of physical and psychological development and the principles of performance growth in a given sport. Dovalil et al. (2009) argue that it is essential to build a healthy attitude to the sport.

Differences in skill acquisition, anatomy, physiology, or performance among individuals during ontogeny have been highlighted in papers such as Buttelmann and Böhm (2014), Farber and Petrenko (2011), and Peric (2008).

The disparities in individual development associated with skill acquisition and motor skills during ontogeny provide a rationale for the development of an appropriate assessment tool and the subsequent establishment of standards concerning player puck control in ice hockey. A standardised assessment tool for different age categories, along with the derived benchmarks, may prove beneficial not only for coaches and methodologists, but also for the players themselves by providing clear insight into their level of skill mastery.

Hence, assessing ice hockey skills appears appropriate, given the substantial differences between individuals. Without applying scientific procedures in the assessment, a simple subjective evaluation may lead to inaccurate conclusions

2.2. A Chapter on Hockey Systematics as a Starting Point for Assessment Tool Development

Existing methodological materials from hockey associations in the USA, Canada, and Czech Republic indicate that fundamental skills include skating, puck control, shooting, passing, and other skills derived from these. Skating is consistently cited as the first skill to be learnt, and in practical terms, skating is also prioritised in training. It can be said that skating forms the foundation and is also a natural prerequisite for ice hockey.

As has been noted, without a perfect mastery of skating skills, it is impossible to reach the top level. It is recommended to commence teaching puck control immediately after mastering the basics of skating or concurrently with skating instruction (Český hokej, 2018; Hockey Canada, 2018; USA Hockey, 2018). Players who have already acquired basic skating skills should start to learn the basics of puck control (Pavliš et al., 2009). Bukač (2014) states that in the developed hockey world, significant attention is given to technique training. Consistent stick technique training, overseen by a coach, should be integral to a player's training regimen.

Minor discrepancies exist in the systems used by various hockey-playing countries, such as differences in terminology or the timing of skill acquisition initiation, often reflecting variations in age group distribution.

Czech Hockey (2018) lists the game activity of advancing the puck as a superordinate term. Hockey Canada (2018) and USA Hockey (2018) use more also more terms, but the term puck control as an umbrella term for all related skills. However, for the purposes of this thesis, it was considered that all these skills and activities of an individual, such as advancing the puck, as part of puck control.

2.3. Laws of Motor Learning as a Basis for Assessment Tool Development

Play and learning are among the most fundamental human activities. In general, learning can be described as an active and creative process involving the acquisition and development of knowledge, skills, abilities, attitudes, and norms. One specific type of learning is motor (sensorimotor) learning.

Motor learning denotes a continuous process of mastery and refinement of motor skills. These changes occur across various levels, stages, sections, and most commonly, phases. The phases of motor learning are interconnected and build upon each other, although their exact timing is not always clear, and they flow seamlessly into one another. During motor learning, three to four phases can be discerned according to qualitative differences. There are five types of learning in the motor learning process; namely, imitational, instructional, feedback-based, problem-based, and ideomotor, with the first two being the most common in children. However, learning is not always linear. Rather, it follows learning curves, with positive factors pushing the curve upwards and negative factors pushing it downwards. The evolving dynamics of learning have been highlighted, for example, by Ghorbani & Bund (2017) and Kalinski et al. (2016). Motor learning leads

to the development of skills. The basic characteristics of a learnt skill include quality, speed, economy, and method of execution.

The variable dynamics of individual learning and acquisition of ice hockey skills are additional reasons to develop an assessment tool in ice hockey, create expert diagnostic scales, and define standards based on these scales.

2.4. Theoretical Background for Data Processing Diagnostic Tool Development

Item response theory (IRT) has been used in the literature to establish standards (Primi et al., 2016). This theory, rooted in statistical-probabilistic relationships.

In kinanthropology, various types of tests and rating scales are employed to assess skills. For their development, cross-disciplinary literature widely recommends the Rasch model (Anshel et al., 2009; Hecimovich et al., 2014) and Guttman scale constructions (Bertoli-Barsotti & Bacci, 2014).

According to Čepička (2003), the Rasch model is gaining popularity in the assessment of motor skills.

The Rasch model was developed for analysing binary data. Based on the Rasch model, we aim to estimate the relationship between the probability of correctly answering an item and the requisite level of ability to complete said item. When item difficulty aligns with the respondent's latent trait level, the probability of completing the item is fifty percent. These properties of the Rasch model are instrumental in constructing an assessment tool. When we know the difficulty of the items, we can organise them in the form of a Guttman scale. The literature also uses the term ideal or perfect scale (Gothwal et al., 2009; Davis-Stober et al., 2015). All these labels are based on the premise that completing an item of a certain difficulty on a given scale should imply automatic completion of all items of lower difficulty. Mokken scaling techniques are a useful tool for developing and constructing unidimensional tests (Sijtsma et al., 2008). Stochl et al. (2012) add that Mokken techniques can also be used to test the appropriateness of using Rasch analysis.

The main advantage of Mokken scaling techniques is that they do not require the assumption of non-linear behaviour of the response probability (Sijtsma & Molenaar, 2002).

The use of Guttman scale and Rasch and Mokken analysis appears to be a suitable means of developing an assessment tool for the evaluation of puck control skills in ice hockey.

3. MATERIALS AND METHODS

This section was structured around the objectives of the theoretical-empirical work concerning the creation of an assessment tool for ice hockey players and the subsequent establishment of standards. The research was divided into three phases.

In the zero phase of the research, a team of four to five experts in the field of ice hockey was selected to assist in the development of the research formula to conduct a systematic review of the literature to verify the absence of a diagnostic tool in ice hockey.

In the first stage, expert "makers" were selected to help compile items for the item bank, and an item bank containing 60-100 items was created. Expert "calibration raters" and calibration[1] participants were selected, and the items of the diagnostic tool was calibrated. Following the calibration, Rasch and Mokken analyses and item selection[2] were carried out, the properties of the scale were assessed, and the items were ranked to form a Guttman scale.

In the second stage expert “test raters” were selected for assessing the players. Also, the players for testing were selected and the test was carried out. The data obtained from the tests underwent statistical processing and standards for puck control in ice hockey for players aged between 6-15 years in Czech Republic were established.

3.1. Aims and Tasks

The objectives of the thesis are outlined as follows:

- The first aim of the thesis is to develop a tool to assess an individual’s ice hockey skill – puck control.

[1] For the purpose this thesis the term item calibration was used to refer to the collection of data for subsequent analysis (verification) of items using Rasch and Mokken analysis. Following from the item calibration and the results of the analysis, item selection was conducted (see under “Selection” for more details).

[2] After item calibration and analysis of the collected data, item selection was carried out. In each round of selection, all items were evaluated based on the data obtained, and items that did not fit the Rasch model were removed from the item bank. In each subsequent round of selection, only the remaining items were analysed using the same key as in the first round.

- The second aim of the thesis is to utilise the developed assessment tool to establish standards pertaining to puck control among ice hockey players aged between 6 and 15 years in Czech Republic.

3.2. Research Questions

Given the two aims of the thesis, it is appropriate to formulate the research question in two parts.

- The first question is: Which puck control skills should ice hockey players have mastered by the age of fifteen?
- The second question is: What are the standards of puck control in each age category of ice hockey?

3.3. Hypotheses

The hypotheses are formulated as follows:

- H1: The most difficult item will not fit the selected model in the developed assessment tool.
- H2: The second most difficult item will not fit the selected model in the developed assessment tool.

3.4. Research Set

Based on predefined conditions, it was necessary to select a group of 4-5 experts in the field of ice hockey to find suitable keywords for systematic review, 4-5 expert "makers" who helped to create an item bank for the diagnostic tool, and a group of approximately 80 expert "calibration raters" and "test raters" for calibration and testing, and a group of 400-600 participants for both calibration (100-200) and testing (200-400).

3.4.1. Statistical Methods

An assessment tool was developed to test selected participants. The Guttman scale, Rasch analysis and Mokken analysis were employed to create an assessment tool for diagnosing ice hockey players in terms of puck control abilities. Validity and reliability of the scale was assessed and item difficulty was calculated. Content validity was determined and reliability was assessed using Cronbach's alpha and Mokken analysis. When defining standards, the mean, median, and standard deviation was calculated.

4. RESULTS

4.1. Systematic Review

The systematic literature review confirmed that no standardised assessment tool was available to assess the ice hockey skill relevant to our research; that is, puck control. Following the literature review and comments of expert "makers", an item bank of 74 dichotomous/binary items[3] was created.

Once the item bank was finished, the items were calibrated. Expert raters and participants were selected and the items were calibrated. A total of 357 participants were selected for calibration and all 74 items[4] were tested. The participants were rated on a yes/no (pass/fail) scale, and their results were then processed and prepared for item validation. Once the item bank items had been calibrated, the data were collected, analysed, and scored using Rasch analysis and Mokken analysis.

During the individual selection rounds, items that did not fit were gradually eliminated from the item bank based on the outfit and infit mean square values. In the first round of selection, 23 items were eliminated from the item bank; in the second round, 9 items were eliminated; in the third round, 6 items were eliminated; in the fourth round, 3 items were eliminated; in the fifth round, 7 items were eliminated; and in the sixth round, 2 items were eliminated. In the seventh round of selection, all items (24) appeared to fit the Rasch model, and no further items needed to be eliminated from the item bank. Although during the selection process, items that did not fit the Rasch model were discarded, the scores of the two most difficult items in the item bank were taken into account, and the items were retained in the bank. This was a deliberate decision despite a slight violation of the rules for removing items from the item bank when they did not fit within the required interval. After seven rounds of selection, item selection, and analysis, 17 items were finally selected for the assessment tool. These were items 3, 7, 10, 12, 20, 39, 40, 49, 52, 60, 61, 62, 63, 67, 68, 71, and 72. Item difficulty values ranged from 2.90 logits for easiest item

[3] Henceforth only the numerical designation of the item was used, without further reference to the full verbal name or explanation of the content or principle of the item. The naming, as well as correct and incorrect designations of items are given in the Appendices. The numerical designation of the items is random and has no influence on the difficulty or the order of the items during their execution. The numerical codes have been introduced mostly for an immediate control over the number of items and their easy identification during the research. In practice, in the assessment tool that has been created, the numbers no longer serve to describe the items but to indicate the order in which the items must be administered to the players in order to assess the player's latent trait.

[4] A list of all the items can be found in Appendix 4 – Items.

(item 3) to −4.16 logits for the most difficult item (item 61). All 17 items included in the assessment tool were analysed, as in the previous rounds. The standard error of the estimate and the lower and upper confidence intervals were also calculated. Goodness-of-fit and Pearson's chi-square were used to check the fit of the model. Its value was $p = 1$. The data fit the model. The descriptive statistics of the model were evaluated, and 92.4% of the items fell within the {−4, 4} logits interval, corresponding to almost 16 out of 17 items.

Again, the item and person fit was calculated. As well as the plot of the item characteristic curves, the item information curve and test information function which shows the dependence of the test information function on the standard error. The reliability of the test, calculated using Cronbach's alpha, was 0.929 with a confidence interval of {0.918; 0.939}. Validity was assessed by expert judgement, as in all previous rounds. Mokken analysis was carried out to confirm the reliability values. The unidimensionality of the test was calculated as $H = 0.691$ with a standard error of $SE = 0.021$. It was verified that there was no monotonicity violation. The reliability assessment showed that there were no violations of invariant item ordering too. The test value was $H = 0.566$, and the reliability calculated using Mokken analysis was $rho = 0.94$.

Even after removing more than three-quarters of the items from the item bank and selecting items, the items remaining in the assessment tool covered the full range of difficulty. This is particularly the case in the {−3, 2} logits interval, where most of the respondents lie according to the latent trait.

The outfit and infit mean square scores were checked, for which a range of 0.5 to 1.5 was set, except for items 60 and 61. Only item 60 did not fit into this interval. The fact that the most difficult item 61 fit the selected model refuted the first hypothesis, H1: The most difficult item will not fit the selected model in the created assessment tool. The second hypothesis, H2: The second most difficult item (item 60) did not fit the selected model in the created assessment tool, was confirmed, as the second most difficult item indeed did not fit the selected model according to the values obtained. However, the item was retained in the tool. The p-values were also verified with the condition of values higher than 0.05.

Finally, the Guttman error was analysed, which showed that 2 out of 3 (almost 250 out of 357) participants have no Guttman error. The high resistance is also confirmed by the calculated index $H = 0.691$ and the standard error 0.021.

Rasch model was used to create an assessment tool for puck control skills in ice hockey. This tool, which includes 17 items, was compiled into a Guttman scale based on the difficulty of the items and was further used to set standards for two-year cohorts in ice hockey in the Czech Republic. Thus, the first aim of this thesis, the creation of a tool for assessing an individual's skill in ice hockey, was achieved.

Once the assessment tool had been developed, data collection was undertaken to establish standards. Data were collected from as many as 1 102 participants aged 6–15 years, with the different age cohorts equally represented with approximately 200 participants per cohort.

For the 6–7-year cohort (U7), data were obtained from 180 participants. On average, players in this category in the Czech Republic completed 0.96 items, the median was 0 items, and the standard deviation was 1.25 items.

For the 8–9-year cohort (U9), data were obtained from 253 participants. On average, players in this category in the Czech Republic completed 2.15 items, with a median of 2 and a standard deviation of 1.9 items.

For the 10–11-year cohort (U11), data were collected from 266 participants. On average, players in this category in the Czech Republic completed 4.56 items, the median was 4, and the standard deviation was 3.24 items.

For the cohort aged 12–13 years (U13), data were gathered from 221 participants. On average, players in this category in the Czech Republic completed 5.73 items, with a median of 6 and a standard deviation of 3.3.

For the 14–15-year cohort (U15), data were obtained from 182 participants. On average, players in this category in the Czech Republic completed 8.2 items; the median was 9, and the standard deviation was 3.25 items. This has set the standards for the puck control in ice hockey in the Czech Republic. Thus, the second aim of this thesis, to establish standards pertaining to puck control among ice hockey players aged between 6 and 15 years in Czech Republic, was achieved.

5. DISCUSSION

A total of 17 items were selected for our assessment tool. Based on the calibration, 24 items fit the Rasch model. In our opinion, however, it was appropriate to reduce the number of items in view of the subsequent use in practice. A smaller number of items indicating the skill level of the player is more suitable for coaches. In addition, the items retained in the tool cover almost the entire range of difficulty and will therefore optimally reflect the level of the latent trait. Some of the discarded items differed only minimally in difficulty from the selected items, so their discriminative value would be negligible compared to the selected items. During calibration it also became apparent that some eventually discarded items were more demanding in terms of preparation, explanation, and overall test administration than the selected items, while their difficulty was similar to that of the retained items. Therefore, we believe that the number of 17 selected items is appropriate for use in the assessment tool.

The Rasch model was chosen for the development of the assessment tool. It is usually recommended in the literature as the simplest model of item response theory for the development of assessment tools (Čepička, 2003; Dragounova, 2018). Due to the nature of the items created for the item bank, the three-parameter model was rejected at the outset. The items were created in such a way that random guessing was not possible or a successful attempt had to be repeated. The two-parameter model was rejected because the number of items was large enough and because we used items that only had to indicate the participant's level of the latent trait, so that the discriminative ability of the items was not important for our instrument. The decision to use a one-parameter model or a Rasch model was then based on the literature (de Ayala, 2009; Linacre, 2002) and the principles underlying the one-parameter and Rasch models. The one-parameter model is used in situations where the model fits the data. The Rasch model was then used to construct tests, and items that did not fit the model tend to be removed from the bank. However, this requires a sufficiently large item bank. All these requirements have been met. Therefore, in our opinion, the use of the Rasch model in this research is appropriate and its use in similar research can be further recommended.

In the process of calibrating the item bank for the purpose of the assessment tool, items that did not fit the Rasch model were discarded. This was done in the first seven rounds of selection based on the outfit mean square and infit mean square statistics, with values in the range of 0.5 to 1.5 being recommended for the development of the

assessment tool. Scores above 1.5 are unproductive for assessment tool development, and scores below 0.5 are less productive but not degrading (Linacre, 2003). Based on the analysis, items 60 and 61 were found to be the most difficult. However, their outfit and infit mean square values during calibration were always at the lower end of the interval recommended for assessment tool development. Due to their difficulty and subsequent use in the assessment tool, and therefore their practical contribution, it was decided that as long as the items did not have extreme values, they would be retained in the item bank, even though their outfit and infit values were not within the interval. During calibration, items 60 and 61 were at the lower end of the interval, with item 60 more likely to have lower outfit values. We attribute the low outfit mean square values to the fact that only a small proportion of participants were able to complete the items during calibration. Subsequently it was used in this developed assessment tool, with items 60 and 61 being the most difficult and thus providing optimal feedback for coaches and players around the age of 15, when much of the selection of players in their future hockey careers will take place. For these reasons, items 60 and 61 were retained in the item bank and subsequently in the assessment tool, even though their outfit and infit mean square values at certain stages of the selection rounds and the choosing of items for the assessment tool indicated that the items should be discarded. In our opinion, the inclusion of items 60 and 61 in the assessment tool will only have a positive effect in practice, justifying their inclusion in the assessment tool despite their not entirely satisfactory scores due to their high difficulty.

To assess reliability, it was decided to use Cronbach's alpha, as in Martínková and Zvara (2007), Christmann and Van Aelst (2006), and McNeish (2018). However, Sijtsma and Molenaar (2002) report that Cronbach's alpha can be affected, whereas the rho-coefficient calculated using Mokken analysis is almost free from side effects. Which is why it was decided to use both coefficients to assess reliability. As mentioned above, the Mokken analysis also helped us to explain the low test information anomaly. Cronbach's alpha increased over the rounds of selection and elimination of nonconforming items, and the rho-coefficient also increased as the prerequisites for using Mokken analysis were met. The values then approached a value of one for the developed assessment tool. Since the two coefficients used were in agreement based on the analyses carried out, we can conclude that the test is consistent and that both coefficients are suitable for assessing reliability in similar research.

When evaluating the skill of puck control, (i.e., not score) the quality of skating had to be ignored. This was due to the requirement of simplicity of the practical application and also for calibration and data processing purposes. Although we do not doubt that the skating ability is a prerequisite for hockey players, for the purposes of our work, we had to disregard this skill. In ice hockey systematics, skating is singled out as a basic skill, and in recommendations for the training process, skating is always mentioned separately. When practicing individual game skills, skating is described separately or it is automatically taken as a necessity for mastering other skills (Hockey Canada, 2018; USA Hockey, 2024). We also used this assumption in our work, where skating was taken as a necessary prerequisite for mastering (fulfilling) a given item. Instead, we based our evaluation of the players on the fact that, just as the difficulty of the items created increases with the length of time the players are trained in puck control, so too should the players improve their skating skills. Although there are items that involve movement and require skating to complete them, and the maximum possible intensity of movement is required, it is only the skill of puck control that is assessed. Skating technique, quality, or fluidity of movement is not assessed. We also assumed that the assessment tool was designed for players aged 6 to 15 and that their skills would naturally be different. It can also be argued that especially in today's hockey, speed of movement and acceleration are essential. In this case, the ability to shoot or pass immediately after stick handling around the opponent/defender ("fake-pull") would also have to be taken into account. However, these are other fundamental skills, and it is necessary to prepare tests specifically for them and then evaluate whether it is possible to create a test for all the skills at the same time. In addition, the definition of the cut-off points for the assessment itself would be very difficult and hardly controllable in practice. In practice, then, as we collected data to set the standards, it became clear that when players did not have sufficient skill, they were often unable to complete the items even at a leisurely pace, let alone at maximum intensity. We acknowledge that in the ideal case, it would be useful to combine all skills into a whole, based on game situations in the game. However, in our opinion, it is unrealistic to expect such an assessment tool to be scientifically based, and the question arises as to whether a complete expert assessment in games would not be preferable for such a purpose. It was also not the intention of our research to replace such an assessment.

6. CONCLUSION

In the zero phase of this research, a systematic review of the literature was conducted, which confirmed that no scientifically based assessment tool for ice hockey skills was available in practice or in the literature. After confirming this initial hypothesis, a 74-item puck control skill item bank was created. The items were calibrated and validated using data obtained from 357 participants. In seven rounds of elimination, items that did not fit the Rasch model were gradually removed from the item bank. Thus, the assessment tool contains 17 items.

Using the Guttman scale, the items were ranked from the easiest to the most difficult. This fulfilled the first aim of this work and answered the first research question. At the same time, hypothesis H1 was rejected (i.e., the hypothesis that the most difficult item will not fit the selected model in the created assessment tool), and hypothesis H2 was confirmed (i.e., the hypothesis that the second most difficult item will not fit the selected model in the created assessment tool).

The 17-item assessment tool developed and compiled in this way was used to define standards in terms of puck control for ice hockey players aged 6–15 years in the Czech Republic for cohorts of 6–7 years, 8–9 years, 10–11 years, 12–13 years, and 14–15 years. This also fulfilled the second aim of our thesis and answered the second research question of our thesis (i.e., what are the standards of puck control in each age category of ice hockey).

To define the standards, data were collected from participants in hockey clubs selected from across the Czech Republic. A total of 1 102 participants took part in the tests. Based on the results of our work and the defined standards in puck control in ice hockey for players aged 6–15 years, we can conclude that the player' skills are at an insufficient level, which corresponds to the assessment of coaches not only of youth national teams. Only a very small percentage of players reached at least the level of the last four items. Only a few individuals had the opportunity to try to achieve the last two items. For further use of the tool in practice, it would be useful to secure data from players from other hockey-developed countries in Europe and overseas.

The methods developed to create the assessment tool for puck control in ice hockey can also be utilised to prepare assessment tools for other skills. In our opinion, it can also be used in other sports games.

The results of this work will be made available to methodologists and coaches of the Czech Ice Hockey Association, as well as to coaches in clubs, the International Ice Hockey Federation, and individual national associations and federations. We believe that they will contribute to the development of the training process of players and the education of ice hockey coaches, not only in youth categories.

The results of the work will also help to improve the process of motor learning, obtaining feedback, and streamlining the training process in youth categories and beyond.

7. References

Anshel, M., Weatherby, N., Kang, M., & Watson, T. (2009). Rasch calibration of a unidimensional perfectionism inventory for sport. *Psychology of Sport and Exercise*, *10*(1), 210-216. https://doi.org/10.1016/j.psychsport.2008.07.006

Bae, J., Lim, T., & O'Sullivan, D. (2024). The role of sport coaches on student-athletes' life skills development and transfer: Using multilevel model analysis. *International Journal of Sports Science & Coaching*, *19*(3), 944-955. https://doi.org/10.1177/17479541231217368

Bertoli-Barsotti, L., & Bacci, S. (2014). Identifying Guttman Structures in Incomplete Rasch Datasets. *Communications in Statistics - Theory and Methods*, *43*(3), 470-497. https://doi.org/10.1080/03610926.2012.665552

Bournival, M., Martini, G., Trudeau, F., & Lemoyne, J. (2023). The science and art of testing in ice hockey: a systematic review of twenty years of research. *Frontiers in Sports and Active Living*, *5*. https://doi.org/10.3389/fspor.2023.1252093

Bukač, L. (2014). *Trénink herní přirozenosti* (1 ed.). Grada Publishing.

Buttelmann, D., & Böhm, R. (2014). The Ontogeny of the Motivation That Underlies In-Group Bias. *Psychological Science*, *25*(4), 921-927. https://doi.org/10.1177/0956797613516802

Čepička, L. (2003). The Rasch model in the motor ability testing. *Journal of Human Kinetics*, *10*, 99-106.

Český hokej. (2018). *Ceskyhokej.org*. Retrieved 2018-11-07, from https://www.ceskyhokej.org/treneri/metodicke-materialy

Davis-Stober, C., Doignon, J., & Suck, R. (2015). A Note on the Eigensystem of the Covariance Matrix of Dichotomous Guttman Items. *Frontiers in Psychology*, *6*, 1767. https://doi.org/10.3389/fpsyg.2015.01767

de Ayala, R. (2009). *The theory and practice of item response theory* (1 ed.). The Guilford Press.

Douglas, A., Rotondi, M., Baker, J., Jamnik, V., & Macpherson, A. (2022). A Comparison of On-Ice External Load Measures Between Subelite and Elite Female Ice Hockey Players. *Journal of Strength and Conditioning Research*, *36*(7), 1978-1983. https://doi.org/10.1519/JSC.0000000000003771

Dovalil, J., Choutka, M., Svoboda, B., Hošek, V., Perič, T., Potměšil, J., Vránová, J., & Bunc, V. (2009). *Výkon a trénink ve sportu* (3. vyd). Olympia.

Dragounova, Z. (2018). Development and standardization of a rating scale designed for floorball skills diagnostics of young school-age children. *Baltic Journal of Health and Physical Activity*, *10*(4), 34-48. https://doi.org/10.29359/BJHPA.10.4.03

Farber, D., & Petrenko, N. (2011). Neurophysiological mechanisms of recognition fragmented images by five- to six-year-old children. *Human Physiology*, *37*(6), 645-654.

Ghorbani, S., & Bund, A. (2017). Throwing Skills: Analysis of Movement Phases in Early Motor Learning. *Perceptual & Motor Skills*, *124*(2), 502-514.

Gothwal, V., Wright, T., Lamoureux, E., & Pesudovs, K. (2009). Guttman Scale Analysis of the Distance Vision Scale, *50*(9). https://doi.org/10.1167/iovs.08-3330

Hecimovich, M., Peiffer, J., & Harbaugh, A. (2014). Development and psychometric evaluation of a post exercise exhaustion scale utilising the Rasch measurement model. *Psychology of Sport and Exercise*, *15*(6), 569-579. https://doi.org/10.1016/j.psychsport.2014.06.003

Hockey Canada. (2018). *Hockey Canada Player Development*. Retrieved 2018-12-26, from https://www.hockeycanada.ca/en-ca/hockey-programs/players/essentials

Charbonnet, B., & Conzelmann, A. (2024). Talent development in childhood: Early specialization or sampling? From an either… or… question to a 2 × 2 × 3 question cuboid. *International Journal of Sports Science & Coaching*, *19*(1), 459-475. https://doi.org/10.1177/17479541231197225

Christmann, A., & Van Aelst, S. (2006). Robust estimation of Cronbach's alpha. *Journal of Multivariate Analysis*, *97*(7), 1660-1674. https://doi.org/10.1016/j.jmva.2005.05.012

Kalinski, S., Jelaska, I., & Labrović, A. (2016). Dynamics of learning complex motor skills in preschool children. *Sport Scientific & Practical Aspects*, *13*(2), 13-16.

Kostka, V. (1963). *Útok v ledním hokeji* (1.). Sportovní a turistické nakladatelství, vydavatelství ÚV ČSTV.

Linacre, J. (2003). Rasch power analysis: size vs. significance: standardized chi-square fit statistics. *Rasch Measurement Transactions*, *17*(1), 918.

Linacre, J. (2002). What do Infit and Outfit, Mean-square and Standardized mean?. *Rasch Measurement Transactions*, *16*(2), 878. https://www.rasch.org/rmt/rmt162f.htm

Martini, G., Brunelle, J., Trudeau, F., & Lemoyne, J. (2018). Measuring ice hockey skills in a repeated measures testing context: the effects of fatigue on skating efficiency, passing, agility, and shooting. *The Sport Journal*, *21*, 1-16.

Martinková, P., & Zvára, K. (2007). Reliability in the Rasch model. *Kybernetika*, *43*(3), 315-326.

Matthews, A., & Erickson, K. (2024). Exploring the use of individualized consideration by minor hockey coaches. *International Journal of Sports Science & Coaching*, *19*(1), 77-87. https://doi.org/10.1177/17479541231193302

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412-433. https://doi.org/10.1037/met0000144

Nadeau, L., Godbout, P., & Richard, J. (2008). Assessment of ice hockey performance in real-game conditions. *European Journal of Sport Science*, *8*(6), 379-388. https://doi.org/10.1080/17461390802284456

Pavliš, Z., Perič, T., Heller, J., Janák, V., Jansa, P., & Čáslavská, E. (2003). *Školení trenérů ledního hokeje: vybrané obecné obory* (1. vyd). Český svaz ledního hokeje.

Pavliš, Z., Perič, T., Novák, Z., & Beránek, J. (2009). *Příručka pro trenéry ledního hokeje. I. část* (1 ed.). Český svaz ledního hokeje.

Perez, J., Guilhem, G., & Brocherie, F. (2022). Reliability of the force-velocity-power variables during ice hockey sprint acceleration. *Sports Biomechanics*, *21*(1), 56-70. https://doi.org/10.1080/14763141.2019.1648541

Perič, T. (2008). *Sportovní příprava dětí* (2., dopl. vyd). Grada.

Perič, T., & Dovalil, J. (2010). *Sportovní trénink* (1.). Grada Publishing.

Primi, C., Morsanyi, K., Chiesi, F., Donati, M., & Hamilton, J. (2016). The Development and Testing of a New Version of the Cognitive Reflection Test Applying Item Response Theory (IRT). *Journal of Behavioral Decision Making*, *29*(5), 453-469. https://doi.org/10.1002/bdm.1883

Raudasoja, I., Vasankari, T., Simula, M., Villberg, J., Vähä-Ypyä, H., & Kokko, S. (2024). Physical education orientation and activity among youth ice hockey coaches. *International Journal of Sports Science & Coaching*, *19*(1), 64-76. https://doi.org/10.1177/17479541231225182

Røsten, S., Sæther, S., Aspvik, N., & Bjørndal, C. (2024). Embedded, Embodied, Enculturated, and Enabling Processes: The Identification and Evaluation of Sporting Talent by Ice Hockey Coaches in Norwegian Youth National Teams. *International Sport Coaching Journal*, *11*(1), 113-123. https://doi.org/10.1123/iscj.2022-0058

Sijtsma, K., Emons, W., Bouwmeester, S., Nyklíček, I., & Roorda, L. (2008). Nonparametric IRT analysis of Quality-of-Life Scales and its application to the World Health Organization Quality-of-Life Scale (WHOQOL-Bref). *Quality of Life Research*, *17*(2), 275-290. https://doi.org/10.1007/s11136-007-9281-6

Sijtsma, K., & Molenaar, I. (2002). *Introduction to Nonparametric Item Response Theory.* (5 ed.). SAGE Publications Inc.

Stochl, J., Jones, P., & Croudace, T. (2012). Mokken scale analysis of mental health and well-being questionnaire item responses: a non-parametric IRT method in empirical research for applied health researchers. *BMC Medical Research Methodology*, *12*(1). https://doi.org/10.1186/1471-2288-12-74

Tarter, B., Kirisci, L., Tarter, R., Jamnik, V., Gledhill, N., & McGuire, E. (2017). Use of the Sports Performance Index for Hockey (SPI-H) to Predict NHL Player Value. *International Journal of Performance Analysis in Sport*, *9*(2), 238-244. https://doi.org/10.1080/24748668.2009.11868480

Urquhart, D., Bloom, G., & Loughead, T. (2020). The Development, Articulation, and Implementation of a Coaching Vision of Multiple Championship–Winning University Ice Hockey Coaches. *International Sport Coaching Journal*, *7*(3), 335-346. https://doi.org/10.1123/iscj.2019-0096

USA Hockey. (2018). *American Development Model*. Retrieved 2018-12-26, from https://www.admkids.com/

Vigh-Larsen, J., Beck, J., Daasbjerg, A., Knudsen, C., Kvorning, T., Overgaard, K., Andersen, T., & Mohr, M. (2019). Fitness Characteristics of Elite and Subelite Male Ice Hockey Players: A Cross-Sectional Study. *Journal of Strength and Conditioning Research*, *33*(9), 2352-2360. https://doi.org/10.1519/JSC.0000000000003285

Wiseman, A., Bracken, N., Horton, S., & Weir, P. (2014). The Difficulty of Talent Identification: Inconsistency among Coaches through Skill-Based Assessment of Youth Hockey Players. *International Journal of Sports Science & Coaching*, *9*(3), 447-455. https://doi.org/10.1260/1747-9541.9.3.447