Validating Expected Possession Value Against Expert Coach Assessments of Passing Intelligence in Elite Youth Ice Hockey

Miikka Palvalin

Finnish Sports Institute, Vierumäki, Finland

Abstract. The quantitative evaluation of passing-related game intelligence in ice hockey remains methodologically underdeveloped. This study examined the convergent validity of Expected Possession Value (EPV) based passing metrics against structured expert coach assessments of passing intelligence in elite youth ice hockey. Data were collected from 120 male players across 12 clubs competing in the Finnish U20 SM-sarja during the 2025–26 season. Coach assessments were obtained via a three-item structured rating instrument demonstrating excellent internal consistency (α = .950) and good-to-excellent inter-rater reliability (ICC = .852). EPV-based passing metrics showed moderate convergent validity (r = .40-.44) with coach assessments across the full sample, consistently outperforming traditional passing metrics such as pass completion percentage. Positional analyses revealed substantially stronger associations for defensemen (EPV Passes+ / 60, r = .70) than for forwards (r = .39). The findings support EPV as a theoretically grounded operationalisation of passing intelligence and highlight player position as a meaningful moderator of metric validity.

Keywords: Game intelligence, Expected Possession Value, Passing metrics, Ice hockey, Player evaluation

1 Introduction

The quantitative evaluation of ice hockey players has undergone substantial transformation over the past two decades. Where traditional box-score statistics such as goals, assists, and plus-minus once dominated player assessment, the field has progressively moved toward metrics that capture the quality and context of actions rather than merely their occurrence [1,2]. The most recent advance in this progression is Expected Possession Value (EPV): a class of models that assigns a continuous value to every on-puck action; passes, carries, and turnovers alike based on its effect on the team's probability of scoring [3, 4, 5]. Unlike traditional metrics, EPV is outcome-independent and sensitive to decision quality, making it theoretically well suited to evaluating dimensions of play.

Game intelligence in team sports is broadly understood as the ability to perceive relevant cues, make appropriate decisions under time pressure, and execute those

decisions effectively a perception–decision–execution chain consistently identified as the cognitive basis of expert performance [6, 7, 8]. Passing is the team sport action that most directly engages all three stages simultaneously: the passer must scan available options, evaluate their risk and reward, and deliver the puck accurately to the chosen recipient before the defensive structure closes. Expert players differ systematically from less skilled performers in these perceptual and decision-making stages, using fewer but longer fixations, anticipating outcomes earlier, and generating higher-quality first options [9].

Despite this theoretical alignment, the empirical relationship between automated tracking data and expert human judgement of passing intelligence has not been established. Existing validation work in hockey analytics has benchmarked metrics against objective outcomes; goal rates, team success, player contracts, rather than against independent assessments of specific cognitive behaviours [1]. The closest precedent is Gréhaigne et al. [13], who demonstrated strong agreement between expert observer rankings and structured observational scores in ice hockey; however, that study employed controlled assessment conditions rather than season-long automated tracking data. Coach assessments, when structured and anchored, have been shown to constitute credible criterion measures in sport science contexts [10, 11, 14], yet their potential as a validation instrument for automated hockey metrics has not been exploited.

A further dimension of the validation question concerns playing position. Forwards and defensemen occupy structurally different roles in ice hockey, with different passing responsibilities, movement patterns, and tactical mandates. Radke et al. [12] demonstrated that defensemen and forwards exhibit systematically different passing behaviours in tracking data. Defensemen distributing the puck to forwards significantly more often than to their partners, forwards completing passes through narrower lanes under higher pressure. These structural differences suggest that the relationship between EPV metrics and perceived passing intelligence may not be uniform across positions, a hypothesis that position-stratified validation is necessary to test.

The present study addresses these gaps by examining the convergent validity of EPV-based ice hockey performance metrics against structured multi-rater coach assessments of passing-related game intelligence in a player level. Data were collected from 120 players across 12 clubs competing in the Finnish U20 SM-sarja (Under-20 National League) during the 2025-26 season. Two research questions structure the research paper:

RQ1: To what extent do EPV-based passing metrics correlate with coach expert assessments of passing-related game intelligence?

RQ2: Does the validity of performance metrics vary by player position (forward vs. defensemen)?

2 Related Work

2.1 Game Intelligence and Passing in Team Sports

Game intelligence is widely recognised as a decisive factor in team sport performance, yet it resists simple operationalisation. At its core, the construct refers to a player's capacity to read a dynamic environment, make appropriate decisions under time pressure, and execute those decisions effectively [15, 16]. Research across multiple team sports has converged on a three-stage cognitive chain perception, decision-making, and execution as the functional architecture underlying skilled play [17, 22]

A substantial body of research documents systematic differences between expert and novice performers in the perceptual and decision-making stages. A meta-analysis by Mann et al. [6] covering 42 studies and 388 effect sizes found that experts outperform non-experts in response accuracy and reaction time, and that they employ fewer, longer-duration visual fixations, a pattern indicative of more efficient extraction of task-relevant information. Ward and Williams [8] demonstrated that elite soccer players use situational probability cues to anticipate passing recipients more accurately than sub-elite players, with the advantage increasing with game complexity.

Passing is the team-sport action that most directly engages all three stages of the cognitive chain simultaneously. The passer must survey available options (perception), weigh their strategic value and risk (decision-making), and deliver the puck accurately to the chosen recipient (execution). Johnson and Raab [9] showed that professional handball playmakers generate action options sequentially and that the first generated option is typically of highest quality, supporting a fast heuristic model of expert decision-making in time-pressured situations. A recent multi-sport interview study including elite ice hockey players found that top performers use visual perception both to anticipate upcoming situations and to verify whether play is unfolding as expected before committing to an action [7]; a dual perceptual function consistent with the scanning-then-deciding sequence.

2.2 Measuring Passing Intelligence: From Coach Assessment to EPV

Coach Assessment as Expert Judgement. Expert coach assessments represent the historically earliest and still most commonly used method for evaluating player-level game intelligence in applied sport settings. Structured rating instruments that direct coach attention to specific behaviours have been shown to produce reliable and valid assessments [23]. Sieghartsleitner et al. [10] found that subjective coach ratings demonstrated acceptable concurrent validity against objective motor and tactical performance tests in soccer, and that coach assessments contributed unique predictive variance for future performance beyond what objective tests alone captured. McCormack et al. [14] similarly concluded that subjective and objective approaches provide complementary information when profiling athletes. More broadly, the question of how subjective expert

judgement relates to objective performance measurement is a longstanding methodological concern in sport science [19].

The validity of coach assessment is, however, moderated by several factors. Inter-rater reliability varies substantially depending on how clearly the rated construct is defined and whether raters share a common reference frame. A study of coach ratings of athlete mental toughness found good inter-rater agreement when a structured instrument was used [18], while studies using unstructured global ratings report considerably lower agreement. Coach assessments are also susceptible to the halo effect – whereby an overall impression of a player disproportionately influences ratings of specific attributes [11].

Traditional Volume and Accuracy Metrics. Traditional hockey statistics approach passing performance through two broad categories: volume measures, which count passes per unit of playing time, and accuracy measures, primarily completion percentage (Pass%). Their limitations for assessing passing intelligence are well documented. Volume metrics do not distinguish between passes that advance play into high-danger areas and passes that merely retain possession in low-risk locations. Completion percentage conflates decision quality with technical execution: a low-risk pass to a nearby team-mate is easily completed, while a high-risk pass that breaks defensive structure carries a lower expected completion rate even if it is the strategically superior choice. Neither metric is sensitive to the context in which a pass is made the defensive pressure faced, the zone of the ice, or the scoring threat generated. The core problem is that traditional metrics are offensively biased and event-sparse, providing limited signal about the decision-making processes that precede shots and goals [1, 12].

Expected Possession Value. Expected Possession Value (EPV) addresses these limitations by assigning a continuous value to every action based on its effect on the team's probability of scoring. The framework was formalised by Cervone et al. [5] for basketball and subsequently extended to soccer [21] and ice hockey. Murto [3, 4] developed a real-time EPV model for ice hockey using synchronised event and positional tracking data from 872 Finnish Liiga games, implemented through the Wisehockey platform. The model frames possession value as a Markov decision process and defines the reward function as whether either team scores within ten seconds of an action. This model underpins the EPV-derived passing variables used in the present study. Radke et al. [12] confirmed, using NHL tracking data, that EPV-based passing metrics capture dimensions of passing performance not reflected in traditional statistics, and documented systematic positional differences between forwards and defensemen.

3 Methodology

3.1 Participants and Sampling

The study sample comprised 120 male ice hockey players drawn from 12 clubs competing in the Finnish U20 SM-sarja (U20 National League) during the 2025-26 regular season. The U20 SM-sarja is the highest tier of under-20 ice hockey competition in Finland, and the participating clubs represent the development programmes of Finland's top professional organisations. Of the 18 clubs competing in the league, 12 agreed to participate in the coach assessment component of the study, yielding a participation rate of 67%. From each participating club, the six forwards and four defensemen with the highest total full strength ice time in the 2025-26 regular season were included, producing a sample of 72 forwards and 48 defensemen.

3.2 Coach Assessment Instrument

Passing-related game intelligence was assessed using a structured coach rating instrument administered an online survey platform in March 2026, after the completion of the regular season. Each participating club received a survey containing three items assessing different dimensions of passing-related game intelligence, which coaches were asked to rate for each of their six forwards and four defensemen listed in the survey. Players were presented in randomised order within each survey to minimise sequential rating effects, whereby a coach's evaluation of one player might influence their rating of the subsequent player. To enhance the reliability of assessments, each club was asked to provide ratings from a minimum of two coaches.

These items were developed to correspond to the three stages of the perception–decision–execution chain identified in the game intelligence literature [6,7]. Each item (Table 1) was rated on a seven-point scale anchored at the extremes and midpoint: 1 = bottom 10% of the league, 4 = league average, 7 = top 10% of the league. This criterion-referenced anchoring was designed to encourage coaches to use the full range of the scale and to rate players relative to the league as a whole rather than relative to their own roster.

Table 1. Descriptive statistics and reliability of the coaches' assessment scale

Item / Scale	Item wording	*M*	*SD*	α if dropped
Perception	The player perceives and identifies multiple passing options.	4.24	1.18	0.906
Decision-making	The player produces decisive or game-breaking passes.	4.30	1.14	0.947
Execution	The player completes passes successfully.	4.26	1.10	0.926

Total (mean of 3 items)	4.27	1.09	α = .950, ω = .951

Note. N = 120. Items rated on a 1-7 scale. α if dropped = Cronbach's alpha with that item removed. Total scale reliability: Cronbach's α = .950; McDonald's ω = .951.

Prior to the main convergent validity analyses, the psychometric properties of the three-item coach instrument were examined to verify its suitability as a criterion measure. Internal consistency was assessed using Cronbach's α and McDonald's ω. All psychometric analyses were conducted in Jamovi (version 2.7) [24]. Internal consistency of the three-item composite was excellent (Cronbach's α = .950, McDonald's ω = .951). Inter-rater reliability, assessed via intraclass correlation (ICC2k, two-way random, average measures), was good-to-excellent (ICC = .852, 95% CI: .809-.887), supporting the use of the coach mean composite as a reliable criterion measure. Descriptive statistics (see also table 2) of the survey results and the density of Coaches assessment variable show that the coaches were able to use the scale well with high standard deviation.

3.3 Performance Data

Objective performance data were obtained directly from Wisehockey (Wisesport Ltd.), an automated real-time sports analytics platform that collects synchronised event and positional tracking data from all U20 SM-sarja games. Data covered the full 2025–26 regular season. All analyses were restricted to 5-on-5 (full-strength) situations and all per-player metrics were normalised to per-60-minutes of 5v5 ice time to ensure comparability across players who differed in total playing time. Across the analytic sample, 5v5 ice time ranged from 6 to 13 hours, with a mean of approximately 10 hours for defensemen and 8,5 hours for forwards.

EPV metrics are derived from the Wisehockey tracking based possession value model [3, 4], which frames puck possession as a Markov decision process and assigns a continuous value to each on-puck action based on its effect on the team's probability of scoring within the following ten seconds. Passes are weighted by the spatial context in which they occur, with passes progressing the puck into the offensive zone and high-danger areas receiving greater positive weight than lateral or retreating passes. EPV related variables in this study are EPV PassTotal / 60, which sums all positive and negative EPV value from each individual pass and is normalized for 60 minutes of 5v5 ice time. EPV Passes+ / 60 sums only EPV from all the passes that has had positive impact on EPV and EPV Passes- / 60 all the negative impacts including missed passes. EPV CarryTotal / 60 which sums all positive and negative EPV value from each individual puck possession that player has had. All variables are based on algorithms provided by Wisehockey.

Table 2. Descriptive statistics for coaches' assessment and game performance variables

Variable	*N*	*M*	*SD*	*Skw*	*Min*	*Max*
Coaches' assessment	120	4.27	1.09	-0.11	1.44	7.00
Passing variables						
EPV PassTotal / 60	120	0.215	0.127	0.972	-0.005	0.676
EPV Passes+ / 60	120	0.454	0.147	0.977	0.143	1.001
EPV Passes- / 60	120	*-0.239*	0.063	*-0.502*	*-0.475*	*-0.088*
Passes / 60	120	53.61	12.72	0.208	27.57	84.59
Pass%	120	0.807	0.038	-0.647	0.690	0.870
Puck possession variables						
EPV CarryTotal / 60	120	0.194	0.115	1.007	-0.008	0.626
PuckControl / 60	120	2.635	0.717	0.709	1.245	5.047
Scoring threat variables						
xG / 60	120	0.551	0.342	0.684	0.095	1.514
Shots / 60	120	9.60	3.241	1.048	3.30	20.50
xG%	120	0.506	0.066	-0.514	0.306	0.673
xGFor / 60	120	2.678	0.481	-0.334	1.332	3.736
xGAgainst / 60	120	2.604	0.441	1.263	1.754	4.074

The primary analyses examined Pearson r correlations between each of the 12 objective performance variables and the coach composite, conducted separately for three analytical groups: all players (n = 120), all defensemen (n = 48), all forwards (n = 72). Statistical significance was set at $p < .05$ throughout. The magnitude of correlations was interpreted using the benchmarks proposed by Hopkins et al. [20]: r = .10 small, r = .30 moderate, r = .50 large, and r = .70 very large. All correlation analyses were conducted in Jamovi [24].

4 Results

4.1 Overview

Table 3 presents Pearson correlations between each objective performance variable and the coaches' assessment mean (three-item composite) across all three analytical groups. Bold values indicate significance. Analyses were conducted separately for all players, all defensemen, all forwards.

Table 3. Correlations between coaches' assessment and game performance variables across player groups.

Variable	Full sample n = 120		Forwards n = 72		Defensemen n = 48	
	r	***p***	***r***	***p***	***r***	***p***
Passing variables						
EPV PassTotal / 60	.398***	< .001	.332**	0.004	.628***	< .001
EPV Passes+ / 60	.437***	< .001	.388***	< .001	.701***	< .001
EPV Passes- / 60	*-.218**	0.017	*-.265**	0.025	*-.180*	0.22
Passes / 60	.372***	< .001	.319**	0.006	.651***	< .001
Pass%	.234*	0.01	0.13	0.276	.430**	0.002
Puck possession variables						
EPV CarryTotal / 60	.249**	0.006	0.161	0.176	.434**	0.002
PuckControl / 60	.351***	< .001	.244*	0.039	.526***	< .001
Scoring threat variables						
xG / 60	0.136	0.139	.316**	0.007	.393**	0.006
Shots / 60	0.08	0.385	0.14	0.24	*-.004*	0.978
xG%	.391***	< .001	.427***	< .001	.352*	0.014
xGfor / 60	.340***	< .001	.376**	0.001	.318*	0.028
xGAgainst / 60	*-.231**	0.011	*-.242**	0.04	*-.218*	0.137

*Pearson's r. Italics indicate negative correlations. * p < .05, ** p < .01, *** p < .001.*

4.2 RQ1: Convergent Validity Across the Full Sample

Across all 120 players, EPV-based passing metrics showed the strongest associations with coach-assessed passing intelligence. EPV Passes+ / 60 emerged as the leading predictor (r = .44), followed by EPV PassTotal / 60 (r = .40), which is considered as moderate correalation [20]. Notably, both outperformed traditional measures: Passes / 60 (r = .37) and Pass% (r = .23), indicating that the quality-sensitive properties of EPV metrics capture dimensions of passing behaviour that volume and accuracy measures do not fully reflect. The negative association of EPV Passes− / 60 (r = −.22) further demonstrates that EPV metrics provide discriminative information in both directions; identifying not only value-generating but also value-reducing passing decisions; a distinction that traditional metrics such as Pass% cannot make.

Secondary associations were observed for puck possession and scoring-threat variables. PuckControl / 60 (r = .35), xG% (r = .39), and xGfor / 60 (r = .34) all reached significance, suggesting that passing intelligence as assessed by coaches also manifests

in broader offensive contributions. Shots / 60 showed no meaningful association (r = .08, ns), confirming that shot volume is orthogonal to the construct regardless of metric type.

The results support RQ1 affirmatively: EPV-based passing metrics correlate meaningfully with coach assessments of passing-related game intelligence and consistently outperform traditional passing metrics, with EPV Passes+ / 60 representing the strongest single predictor in the full sample.

4.3 RQ2: Position-Specific Validity

Defensemen. Across all 48 defensemen, EPV-based passing metrics showed large associations with coach-assessed passing intelligence: EPV Passes+ / 60 (r = .70) and EPV PassTotal / 60 (r = .63) were the strongest predictors which is considered strong correlation [20]. EPV Passes- / 60 did not reach significance among defensemen (r = -.18), suggesting that value-reducing passing decisions are less diagnostically relevant for this position than the positive value generated.

Secondary associations were observed for puck possession and scoring-threat variables: PuckControl / 60 (r = .53), EPV CarryTotal / 60 (r = .43), xG / 60 (r = .39), and xGfor / 60 (r = .32) all reached significance, while Shots / 60 showed no association.

Forwards. For the 72 forwards, EPV-based passing metrics remained significant but showed weaker associations than among defensemen. EPV Passes+ / 60 (r = .39) and EPV PassTotal / 60 (r = .33) were the strongest passing predictors with moderate strength of correlation. EPV Passes- / 60 reached significance (r = -.27), suggesting that value-reducing passing decisions are more diagnostically relevant for forwards than for defensemen. Among secondary variables, scoring-threat metrics played a more prominent role for forwards: xG% (r = .43) and xGfor / 60 (r = .38) were among the strongest predictors overall, reflecting the position-specific emphasis on offensive production in how coaches evaluate forward play.

Across both positions, EPV-based passing metrics consistently outperformed their traditional counterparts. EPV Passes+ / 60 outperformed Passes / 60 among both defensemen (r = .70 vs. r = .65) and forwards (r = .39 vs. r = .32), and Pass% failed to reach significance among forwards (r = .13, ns) while EPV passing metrics remained significant. EPV Passes+ / 60 was the passing metric to reach the most significance across all three analytical groups, supporting its treatment as the most robust single-metric operationalisation of passing-related game intelligence in this dataset.

The results support RQ2 affirmatively and reveal a clear positional asymmetry. EPV-based passing metrics are substantially more strongly associated with coach-assessed passing intelligence for defensemen than for forwards, and consistently outperform traditional metrics across both positions.

5 Discussion

5.1 RQ1: EPV Metrics Converge with Expert Coach Assessments

The convergence between coach assessments and EPV metrics observed here is broadly consistent with prior work [10, 14] examining the relationship between expert judgement and objective performance data in sport. The present results extend this line of evidence to automated tracking-derived metrics in ice hockey, and further specify the conditions under which convergence is strongest: it is the quality-sensitive, context-weighted EPV metrics that align most closely with what coaches observe and evaluate. This suggests that EPV metrics are not merely a more sophisticated computational tool but a closer empirical approximation of the cognitive dimensions that expert observers attend to when assessing passing intelligence.

EPV Passes+ / 60 emerged as the strongest single predictor (r = .44), followed by EPV PassTotal / 60 (r = .40), while their traditional counterparts, Passes / 60 (r = .37) and Pass% (r = .23), showed systematically weaker associations. Correlation around 0,4 level is considered as moderate [20] in behavioral sciences. This ordering is theoretically coherent: EPV metrics weight passing decisions by the spatial context and scoring-threat consequences they produce, whereas volume and accuracy metrics treat all completed passes as equivalent regardless of where and under what pressure they occur. Coaches, as the results indicate, reward progressive decision quality rather than frequency; precisely the distinction the game intelligence literature identifies as central to expert passing behaviour [6, 9].

5.2 RQ2: Position-Specific Validity and Its Theoretical Interpretation

Defensemen. The strength of the defensemen results is theoretically coherent given the structural properties of the position. Radke et al. [12] documented that defensemen distribute the puck to forwards significantly more often than to their defensive partners, and that this directional consistency makes their passing behaviour more systematic and observable. When the primary passing role is to initiate offensive sequences from the defensive zone, the progressive value of those decisions; precisely what EPV Passes+ corresponds with is the central dimension of performance. EPV Passes- / 60 did not reach significance among defensemen (r = -.18), suggesting that value-reducing decisions are less diagnostically relevant for this position than the positive value generated a finding consistent with the directional asymmetry of the defensemen role.

Forwards. For forwards, convergent validity was present but consistently weaker, and the predictor profile shifted meaningfully toward scoring-threat metrics. xG% (r = .43) and xGfor / 60 (r = .38) were among the strongest predictors alongside EPV Passes+ / 60 (r = .39). However, xGfor / 60 and xG% are team-context metrics that partly reflect offensive zone time and teammate quality rather than individual passing decisions

specifically. One interpretation is that coaches evaluate forward passing intelligence partly through its immediate offensive outcomes: forwards operate closer to the opposition goal, where the connection between a good pass and a scoring chance is direct and visible, while defensemen's passing quality is more often expressed through play-building sequences whose value materialises several events later. Another reasoning could be that it is more difficult for the coaches to evaluate passing for the forwards for some reason. The forward associations with these variables should therefore be interpreted cautiously. High correlations with other than pass related metrics could potentially be also a cause of Halo-effect like Jekauc & Fritsch [11] have found. Disentangling these contributions would require teammate- and context-adjusted versions of the scoring-threat variables, which the present data do not provide. Notably, EPV Passes− / 60 reached significance among forwards (r = -.27) but not defensemen, indicating that value-reducing decisions are more diagnostically relevant for forward evaluation consistent with the higher decision complexity and risk-reward trade-offs forwards face in offensive zone situations [7,9].

The consistent outperformance of EPV over traditional metrics holds across both positions. Pass% failed to reach significance among forwards (r = .13, ns) while EPV passing metrics remained significant, and EPV Passes+ / 60 outperformed Passes / 60 among both defensemen (r = .70 vs. r = .65) and forwards (r = .39 vs. r = .32). The results thus support RQ2 affirmatively: EPV-based metrics correspond more closely to coach assessments of passing intelligence across both positions, and the positional asymmetry in the magnitude of that correspondence reflects genuine structural differences in how passing intelligence is expressed and evaluated at each position.

6 Conclusion

The present findings carry several theoretical implications for both the game intelligence and hockey analytics literatures. First, the convergent validity between EPV-based passing metrics and structured coach assessments provides at least moderate empirical support for the theoretical proposition that EPV captures decision-quality dimensions of passing behaviour. Second, the positional asymmetry in convergent validity extends current theoretical understanding of position-specific performance structures in ice hockey. Moderate correlations with variables non-related to passing were surprising and require follow up studies to investigate their relationship with game intelligence.

The results offer concrete guidance for coaches and performance staff working in elite youth ice hockey. EPV Passes+ / 60 emerged as the strongest single-metric predictor of coach-assessed passing intelligence across both positions, suggesting that this variable provides a reliable and efficient complement to subjective observation in player evaluation. Coaches may use EPV-based passing metrics to identify players whose decision quality is systematically high or low relative to peers, supplementing what is visible in game observation with context-weighted quantitative evidence. The positional asymmetry in the findings has direct practical relevance: EPV metrics are

particularly informative for evaluating defensemen, where the play-building role produces a consistent and interpretable passing signal. For forwards, EPV passing metrics capture moderately the passing intelligence with the coaches judgement, but the question remains open if the bias is in the judgement or in the algorithm.

Several limitations warrant acknowledgement and point toward productive future directions. First, the correlational design of the study cannot establish whether EPV metrics causally reflect passing intelligence, especially for forwards. Future work should examine what causes the difference in correlation level between defenders and forwards and find answer to the question whether it is in the coaches bias or in the algorithm. Second, while the coach assessment instrument demonstrated excellent internal consistency and good-to-excellent inter-rater reliability at the composite level, inter-rater reliability at the individual item level was moderate, suggesting that the three dimensions of passing intelligence are assessed more consistently when aggregated than when evaluated separately. Test-retest reliability and sensitivity to halo effects remain uncharacterised, and discriminant validity analyses would further strengthen the evidence base for the instrument's use as a criterion measure in future work. Third, coach assessments in this study are intra-team ratings, meaning each player was evaluated by coaches from their own club. Future work should examine whether EPV metric validity holds when players are assessed on a common scale by external scouts or standardised scouting instruments. Finally, the data is collected from elite youth male players and it remains unclear if the results can be generalized to other player groups.

References

1. Nandakumar, N., Jensen, S.T.: Historical perspectives and current directions in hockey analytics. Annual Review of Statistics and Its Application 6, 19–36 (2019).
2. Macdonald, B.: An expected goals model for evaluating NHL teams and players. In: Proceedings of the MIT Sloan Sports Analytics Conference (2012).
3. Murto, F.: Towards a real-time possession value framework in ice hockey. In: Proceedings of the Linköping Hockey Analytics Conference (LINHAC 2023), Linköping Electronic Conference Proceedings 201 (2023).
4. Murto, F.: Real-time expected possession value estimation in ice hockey. Master's thesis, Tampere University (2024).
5. Cervone, D., D'Amour, A., Bornn, L., Goldsberry, K.: A multiresolution stochastic process model for predicting basketball possession outcomes. Journal of the American Statistical Association 111(514), 585–599 (2016).
6. Mann, D.T.Y., Williams, A.M., Ward, P., Janelle, C.M.: Perceptual-cognitive expertise in sport: a meta-analysis. Journal of Sport and Exercise Psychology 29(4), 457–478 (2007).
7. Magnaguagno, L., & Beck, D. (2025). Decision-making process in game sports: what do top-level players think of current research?. Frontiers in Sports and Active Living, 7, 1653834.
8. Ward, P., Williams, A.M.: Perceptual and cognitive skill development in soccer: the multidimensional nature of expert performance. Journal of Sport and Exercise Psychology 25(1), 93–111 (2003).

9. Johnson, J.G., Raab, M.: Take the first: option-generation and resulting choices. Organizational Behavior and Human Decision Processes 91(2), 215–229 (2003).
10. Sieghartsleitner, R., Zuber, C., Zibung, M., & Conzelmann, A. (2019). Science or coaches' eye?–Both! Beneficial collaboration of multidimensional measurements and coach assessments for efficient talent selection in elite youth football. Journal of sports science & medicine, 18(1), 32.
11. Jekauc, D., Fritsch, J.: The assessment of the halo-effect in a questionnaire rating dimensions of coaches' competence. International Journal of Sport and Exercise Psychology 23(1), 163–179 (2025).
12. Radke, D., Lu, J., Woloschuk, J., Le, T., Radke, D., Liu, C., Brecht, T.: Analyzing passing metrics in ice hockey using puck and player tracking data. In: Proceedings of LINHAC 2023, Linköping Electronic Conference Proceedings 201 (2023).
13. Gréhaigne, J.F., Richard, J.F., Griffin, L.: The validity and reliability of a performance assessment procedure in ice hockey. Physical Education and Sport Pedagogy 14(1), 65–82 (2008).
14. McCormack, S., Jones, B., Elliott, D., et al.: Coaches' assessment of players' physical performance: subjective and objective measures are needed when profiling players. European Journal of Sport Science 22(4), 557–567 (2022).
15. Furley, P., Memmert, D.: Creativity and working memory capacity in sports: working memory as a limiting factor of successful tactical performance? PLoS ONE 8(7), e68903 (2013).
16. Roca, A., Ford, P.R., McRobert, A.P., Williams, A.M.: Identifying the processes underpinning anticipation and decision-making in a dynamic time-constrained task. Cognitive Processing 14(3), 259–266 (2013).
17. Voigt, L., Friedrich, J., Grove, P., Heinrich, N., Ittlinger, S., Iskra, M., ... & Raab, M. (2023). Advancing judgment and decision-making research in sport psychology by using the body as an informant in embodied choices. Asian Journal of Sport and Exercise Psychology, 3(1), 47-56.
18. Roberts, A. H., Greenwood, D., Humberstone, C., & Raynor, A. J. (2020). Pilot study on the reliability of the coach's eye: identifying talent throughout a 4-day cadet judo camp. Frontiers in Sports and Active Living, 2, 596369.
19. O'Donoghue, P.: Research Methods for Sports Performance Analysis. Routledge, London (2010)
20. Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine & Science in Sports & Exercise, 41*(1), 3–13.
21. Fernández, J., Bornn, L., Cervone, D.: A framework for the fine-grained evaluation of the instantaneous expected value of soccer possessions. Machine Learning 110(6), 1389–1427 (2021).
22. Ashford, M., Abraham, A., & Poolton, J. (2021). Understanding a player's decision-making process in team sports: a systematic review of empirical evidence. Sports, 9(5), 65.
23. Hendry, D. T., Williams, A. M., & Hodges, N. J. (2018). Coach ratings of skills and their relations to practice, play and successful transitions from youth-elite to adult-professional status in soccer. Journal of Sports Sciences, 36(17), 2009-2017.
24. The jamovi project: jamovi, Version 2.7 [Computer Software] (2025). https://www.jamovi.org.