



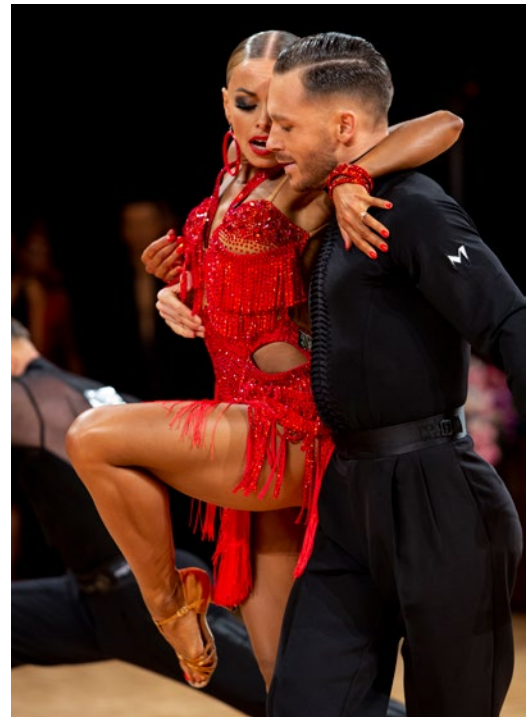
Jerneja Premelč

Sodniška doslednost in zanesljivost ocenjevanja plesnih nastopov v tekmovalnih okoliščinah

Izvleček

Dosledno in nepristransko sojenje je temelj pravičnosti ocenjevanja pri vsakem športu. V estetskih športih je zanesljivost ocenjevanja težje doseči, saj nanj vplivajo tudi drugi subjektivni dejavniki. Študija je proučevala sodniško doslednost in zanesljivost ocenjevanja v standardnih (ST) in latinskoameriških (LA) plesih na svetovnem pokalu WDSF 2024. V raziskavi je sodelovalo 19 najboljših članskih plesnih parov in šest sodnikov, ki so ocenjevali tehnično kakovost, partnerske veščine, gibanje z glasbo in koreografijo s prezentacijo. Rezultati so pokazali dobro zanesljivost ocenjevanja (Cronbach $\alpha = 0,9$; ICC = $0,8-0,9$), pri čemer je bila nekoliko nižja pri tehnični kakovosti in partnerskih veščinah v ST plesih ter pri gibanju z glasbo in koreografiji v LA plesih. Analiza posameznih plesov je pokazala nekoliko nižjo zanesljivost pri počasnem in hitrem fokstrotu v ST plesih ter pri rumbi in džajvu v LA plesih. Med sodniki so bile zaznane individualne razlike v doslednosti ocenjevanja, kar nakazuje na vpliv subjektivnih dejavnikov. Kljub dobri splošni zanesljivosti rezultati opozarjajo na potrebo po nadaljnjem izboljšanju sodniškega sistema, predvsem pri natančnejši opredelitvi kriterijev, usposabljanju sodnikov in razvoju mehanizmov za zmanjševanje pristranskosti. Ugotovitve so pomembne za zagotavljanje večje objektivnosti, pravičnosti in kredibilnosti tekmovanj v športnem plesu.

Ključne besede: standardni plesi, latinskoameriški plesi, zanesljivost, sodniški sistem.



Judging Consistency and Reliability of Dance Performance Evaluation in Competitive Conditions

Abstract

Consistent and impartial judging forms the foundation of fairness in sports evaluation. In aesthetic sports, achieving reliability in judging is particularly challenging, as assessments are also influenced by various subjective factors. This study examined judging consistency and reliability in Standard (ST) and Latin American (LA) dances at the WDSF World Cup 2024. Nineteen top adult dance couples and six judges participated, with evaluations focusing on technical quality, partnering skills, movement to music, and choreography with presentation. The results demonstrated good reliability of judging (Cronbach's $\alpha = 0.9$; ICC = $0.8-0.9$), though slightly lower reliability was observed for technical quality and partnering skills in ST dances, as well as for movement to music and choreography in LA dances. The analysis of individual dances revealed somewhat lower reliability in slow foxtrot and quickstep among ST dances, and in rumba and jive among LA dances. Moreover, individual differences in judges' consistency were noted, indicating the influence of subjective factors. Despite the overall good reliability, the findings underscore the need for further refinement of the judging system, particularly through more precise definition of criteria, regular judge training, and the development of mechanisms to reduce bias. These insights are essential for strengthening objectivity, fairness, and credibility in competitive DanceSport.

Keywords: ballroom dances, latin-american dances, reliability, judging system.

■ UVOD

V standardnih in latinskoameriških plesih se je sodniško ocenjevanje v zadnjem desetletju precej spremenilo in dopolnjevalo. Ocenjevalne kriterije so v pravilniku natančneje opredelili in določili ocenjevalno lestvico. Tako je sodniški sistem postal bolj zanesljiv, manj subjektiven, nadgradnja pa pomeni tudi boljšo povratno informacijo plesalcem o njihovem nastopu (Premelč idr., 2019; Premelč, 2022). Na nižjih ravneh tekmovanja in v mlajših starostnih kategorijah se uporablja star sodniški sistem, ki temelji na razvrščanju plesnih parov. Na najvišjih plesnih tekmovanjih v standardnih (ST) in latinskoameriških (LA) plesih pa sodniki ocenjujejo štiri kriterije, in sicer tehnično kakovost, gibanje z glasbo, partnerske veščine in koreografijo s prezentacijo. Tehnična kakovost se nanaša na natančnost, čistost in obvladovanje plesne tehnike. Gibanje z glasbo meri, kako dobro plesalci interpretirajo in izražajo glasbo skozi gibanje. Partnerske veščine se nanašajo na interakcijo in sodelovanje med plesalcema, pri kriteriju koreografije in predstavitve pa sodniki ocenjujejo ustvarjalnost, strukturo in umetniški vtis. V finalnih nastopih šest sodnikov za posamezni plesni par navede dve oceni na desetstopenjski lestvici. Ena vključuje kriterij tehnične kakovosti (TK) in partnerskih veščin (PV), druga ocena pa vključuje kriterij gibanja z glasbo (GG) in koreografijo s prezentacijo (KP) (WDSF, 2024). Med ST plese spadajo angleški valček, tango, dunajski valček, počasni in hitri fokstrot, med LA plese pa samba, rumba, ča-ča-ča, pasodoble in džajv (jive). Ocenjevalni kriteriji za posamezni ples so sicer enaki, vendar pa so si plesi med seboj različni tako po tehničnih in partnerskih zahtevah, muzikalčnosti in koreografskih značilnosti. V sodniškem pravilniku so tehnične zahteve za vsak ples natančneje določene.

Zanesljivost sodniškega ocenjevanja je izziv ne le v ST in LA plesih, temveč tudi v drugih estetskih športih. V večini teh so se sodniški sistemi spremenili s ciljem izboljšanja ocenjevanja. Več študij je proučevalo zanesljivost sodniškega

ocenjevanja. V športni gimnastiki so te spremembe močno izboljšale zanesljivost ocenjevanja (Atiković idr., 2011; Bučar idr., 2012; Leskošek idr., 2010; Pajek idr., 2013, 2014). Prav tako se je zanesljivost izboljšala pri umetnostnih drsalcih (Lockwood idr., 2005). Premelč idr. (2019) so poročali o slabši zanesljivosti ocenjevanja pri ST in LA plesih v primerjavi z gimnastiko in umetnostnim drsanjem. Poznejša sprememba ocenjevanja z večjim številom sodnikov in manj številnimi kriteriji pa je zanesljivost kljub temu izboljšala (Premelč, 2022). Slabšo zanesljivost so ugotovili tudi pri sojenju hip-hopa, kjer kriteriji ocenjevanja in izvedbe niso dovolj natančno opredeljeni (Sato, 2022). Pri ocenjevanju v estetskih športih je pomembno tudi razumevanje vpliva kognitivnih in socialnih dejavnikov na subjektivne sodniške odločitve. Nanje lahko vplivajo različni dražljaji in dejavniki, na primer odprta povratna informacija, pri čemer se sodniki nezavedno prilagodijo povprečju preostalih ocen (Boen idr., 2008), pozornost sodnika in zaporedje nastopov (Plessner in Schallies, 2005), favoriziranje svojih tekmovalcev in nacionalna pripadnost (Singina, 2021), videz plesalcev (Sato in Hopper, 2021) idr.

Namen študije je bil proučiti sodniško doslednost in zanesljivost ocenjevanja v posameznih plesih. V dozdajšnjih raziskavah je bila zanesljivost ocenjevanja v ST in LA plesih preverjena le na podlagi videoposnetkov, saj je bilo težko zagotoviti zadostno število sodnikov za preverjanje zanesljivosti ocenjevanja posameznih kriterijev. Zaradi spremembe pravil in s tem večjega števila sodnikov je bilo ocenjevanje tokrat mogoče preveriti tudi v tekmovalnem okolju. Ugotovitve lahko pomagajo pri prepoznavanju pomanjkljivosti opredelitve kriterijev za posamezen ples in odpirajo možnosti za izboljšave. Le na podlagi takšnih spoznanj je mogoče uvajati ukrepe, ki bodo izboljšali sodniško doslednost, objektivnost in pravičnost tekmovanj, naj gre za spremembe pravil, dodatno izobraževanje sodnikov ali uporabo novih tehnologij.

■ METODE

Vzorec

V raziskavo je bilo vključenih 19 najboljših članskih plesnih parov in šest sodnikov na svetovnem pokalu WDSF leta 2024. Vsak plesni par je odplesal vseh deset ST in LA plesov. Skupaj so sodniki ocenili 190 plesnih nastopov.

Postopek

Šest sodnikov je na svetovnem pokalu WDSF ocenilo 19 najboljših članskih plesnih parov po novem sodniškem sistemu. Vsak plesni par je pri vsakem sodniku dobil dve oceni, in sicer skupno oceno (O1) za tehnično kakovost (TK) in partnerske veščine (PV) ter skupno oceno (O2) za gibanje z glasbo (GG) in koreografijo s prezentacijo (KP). Ocenjevalna lestvica je bila od 0 do 10 z možnostjo polovične ocene.

Obdelava podatkov

Za analizo rezultatov je bila uporabljena opisna statistika za izračun povprečij ter minimalnih in maksimalnih vrednosti. Za ugotavljanje zanesljivosti je bil izračunan koeficient Cronbach alfa (α) in interklasni koeficient (ICC) za absolutno skladnost. Doslednost ocenjevanja posameznega sodnika je bila prikazana z grafi (boxplot). Statistična značilnost je bila določena pri $p < 0,05$. Za analizo rezultatov je bil uporabljen program SPSS, verzija 29.

REZULTATI

Rezultati v Tabeli 1 prikazujejo dobro zanesljivost ocenjevanja v ST in LA plesih v vseh kriterijih. Nekoliko nižja zanesljivost je znana pri O1 (TK in PV) v ST plesih (ICC = 0,76) in Q2 (GG in KP) v LA plesih (ICC = 0,79).

Pri ST plesih je opaziti, da je zanesljivost slabša v vseh plesih pri O1 (TK in PV) v primerjavi z O2 (GG in KP). Pri večini plesov je zanesljivost dobra, zmerna pa je pri O1 v počasnem in hitrem fokstrotu (ICC = 0,74) (Tabela 2).

Tako kot pri ST je tudi pri LA plesih zanesljivost ocenjevanja dobra, vendar pa je pri LA v nasprotju s ST zanesljivost ocenjevanja boljša pri O1 (TK in PV) v primerjavi z O2 pri vseh plesih, razen rumbe (Tabela 3).

Tabela 1

Zanesljivost ocenjevanja standardnih in latinskoameriških plesov

	Povprečje	Min	Max	Cronbach alfa (α)	ICC	p
ST	7,5	6,7	7,9	0,93	0,89	0,001
ST 1	7,4	6,7	7,9	0,85	0,76	0,001
ST 2	7,6	7,3	7,8	0,87	0,84	0,001
LA	7,5	6,9	7,8	0,92	0,89	0,001
LA 1	7,4	6,9	7,7	0,86	0,81	0,001
LA 2	7,6	7,3	7,8	0,81	0,79	0,001
Skupaj	7,5	6,8	7,8	0,92	0,89	0,001
Skupaj 1	7,4	6,8	7,8	0,85	0,79	0,001
Skupaj 2	7,6	7,3	7,8	0,84	0,82	0,001

Opomba. ST = standardni plesi, LA = latinskoameriški plesi, 1 = ocena 1 (tehnična kakovost in partnerske veščine), 2 = ocena 2 (gibanje z glasbo in koreografija s prezentacijo)

Tabela 2:

Zanesljivost ocenjevanja posameznih standardnih plesov

	Povprečje	Min	Max	Cronbach alfa (α)	ICC	p
AV	7,5	7,0	7,9	0,94	0,91	0,001
AV 1	7,4	7	7,9	0,9	0,80	0,001
AV 2	7,6	7,4	7,9	0,87	0,84	0,001
TG	7,5	6,9	7,9	0,92	0,89	0,001
TG 1	7,4	6,9	7,9	0,84	0,78	0,001
TG 2	7,7	7,4	7,9	0,87	0,84	0,001
DV	7,5	6,7	7,8	0,93	0,90	0,001
DV 1	7,3	6,7	7,9	0,87	0,80	0,001
DV 2	7,7	7,3	7,9	0,87	0,83	0,001
PF	7,5	6,6	8	0,93	0,90	0,001
PF 1	7,4	6,6	8	0,85	0,74	0,001
PF 2	7,6	7,3	8,0	0,86	0,83	0,001
KS	7,5	6,4	8,0	0,94	0,90	0,001
KS 1	7,4	6,4	8,0	0,86	0,74	0,001
KS 2	7,6	7,2	7,9	0,9	0,88	0,001

Opomba. AV = angleški valček, TG = tango, DV = dunajski valček, PF = počasni fokstrot, KS = hitri fokstrot, 1 = ocena 1 (tehnična kakovost in partnerske veščine), 2 = ocena 2 (gibanje z glasbo in koreografija s prezentacijo)

Tabela 3:

Zanesljivost ocenjevanja posameznih latinskoameriških plesov

	Povprečje	Min	Max	Cronbach alfa (α)	ICC	p
SA	7,5	6,9	7,8	0,92	0,89	0,001
SA 1	7,4	6,9	7,8	0,89	0,83	0,001
SA 2	7,6	7,4	7,8	0,80	0,78	0,001
ČČČ	7,5	7,0	7,8	0,92	0,90	0,001
ČČČ 1	7,4	7,0	7,7	0,87	0,82	0,001
ČČČ 2	7,6	7,4	7,8	0,82	0,80	0,001
RU	7,5	6,9	8,0	0,93	0,90	0,001
RU 1	7,4	6,9	7,7	0,85	0,80	0,001
RU 2	7,6	7,4	8,0	0,86	0,84	0,001
PD	7,5	6,7	7,8	0,92	0,90	0,001
PD 1	7,3	6,7	7,6	0,90	0,84	0,001
PD 2	7,6	7,3	7,8	0,81	0,78	0,001
JV	7,5	6,9	7,9	0,90	0,88	0,001
JV 1	7,4	6,9	7,7	0,83	0,79	0,001
JV 2	7,6	7,3	7,9	0,80	0,78	0,001

Opomba. SA = samba, ČČČ = ča-ča-ča, RU = rumba, PD = počasni fokstrot, JV = džajv, 1 = ocena 1 (tehnična kakovost in partnerske veščine), 2 = ocena 2 (gibanje z glasbo in koreografija s prezentacijo)

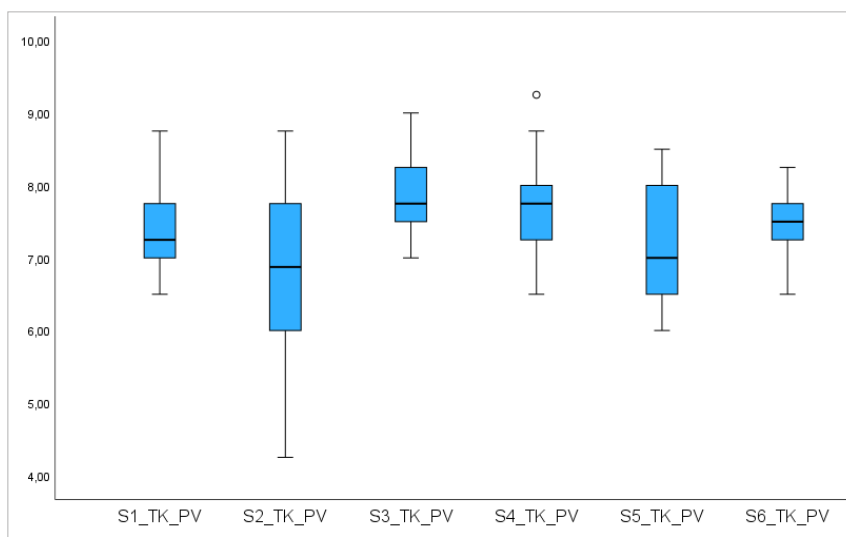
Slika 1 prikazuje, da večina sodnikov TK in PV ocenjuje v območju ocen med 7 in 8, kar kaže na osnovno usklajenost, vendar pa drugi in peti sodnik z zelo širokim razponom ocen zmanjšujeta zanesljivost. Tretji in šesti sodnik sta pri ocenjevanju najbolj dosledna, prvi in četrti sodnik pa imata srednjo konsistentnost.

Slika 2 prikazuje zanesljivost ocenjevanja GG in KP, ki ga večina sodnikov ocenjuje z ocenami med 7,5 in 7,8. Najbolj konsistentno ocenjujejo peti, četrti in šesti sodniki, najmanj konsistentno in sistematično nižje pa tretji nik.

■ RAZPRAVA

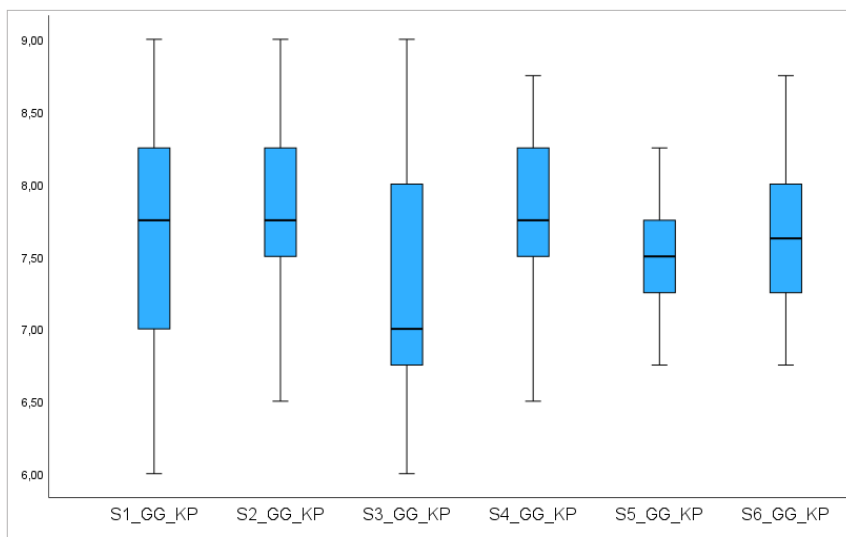
Rezultati naše raziskave kažejo dobro zanesljivost sodniškega ocenjevanja v vseh kriterijih tako v ST kot LA plesih. Nekoliko presenetljiva je ugotovitev, da je pri ST plesih TK in PV nekoliko nižja v primerjavi z LA plesi, saj bi glede na strožje opredeljena tehnična pravila v ST plesih pričakovali ravno obrnjen. ST plesi imajo namreč natančno določeno tehniko (držo, korake, položaj stopal, gibanje po plesišču), zato naj bi sodniki teoretično lažje in enotneje vrednotili ta dva kriterija. Kljub temu so naši podatki pokazali, da so bili sodniki glede ocenjevanja TK in PS skladnejši pri LA plesih. To bi lahko pomenilo, da so morebitne tehnične napake ali razlike med pari v LA plesih očitnejše, medtem ko v ST plesih o niansah tehnike sodniki kljub temu sodijo nekoliko bolj različno. Poleg tega lahko domnevamo, da v ST plesih odločajo manjše podrobnosti (npr. minimalna odstopanja v drži ali gibanju), ki jih vsak sodnik drugače interpretira, kar bi pojasnilo nekoliko nižjo doslednost ocen v primerjavi z LA plesi. Vsekakor ta ugotovitev nakazuje, da jasno določen pravilnik sam po sebi ne zagotavlja popolne usklajenosti sodnikov, saj na ocenjevanje vedno vplivata tudi subjektivna zaznava in interpretacija kriterijev pri posameznih sodnikih.

Na splošno je bila zanesljivost ocenjevanja v naši študiji dobra, kar pomeni, da so sodniki v povprečju dokaj skladno ocenili plesne nastope. Ugotovitev je spodbudna, saj visoka zanesljivost prispeva k večji pravičnosti in kredibilnosti tekmovanj v športnem plesu. Tudi v drugih



Slika 1: Zanesljivost ocenjevanja tehnične kakovosti in partnerskih veščin za posameznega sodnika

Opomba. S_TK_PV = Sodnik_tehnična kakovost in partnerske veščine



Slika 2: Zanesljivost ocenjevanja gibanja z glasbo in koreografije s prezentacijo za posameznega sodnika

Opomba. S_GG_KP = Sodnik_gibanje z glasbo in koreografija s prezentacijo

estetskih športih poročajo o visoki skladnosti sodnikov. V moški športni gimnastiki so izmerili Cronbachov α med 0,92 in 0,99, kar kaže na izjemno konsistentnost ocen (Leskošek idr., 2010). Prav tako so med sodniške korelacije v umetniškem drsanju izredno visoke ($r > 0,90$ za tehnične in umetniške ocene) (Sato, 2022), kar pomeni, da različni sodniki skoraj enako ocenijo posamezen nastop.

Kljub splošno dobri zanesljivosti pa naša analiza razkriva pomembne razlike med

posameznimi sodniki. Ugotovili smo, da nekateri ocenjujejo zelo dosledno, medtem ko drugi pri svojih ocenah variirajo nekoliko bolj. Takšne razlike v sodniški doslednosti so opazili tudi v preteklih raziskavah. Premelč idr. (2019) so pri vrednotenju novega ocenjevalnega sistema v športnem plesu ugotovili, da so bile razlike med ocenami sodnikov precej velike. To kaže, da so se sodniki med seboj v oceni razhajali ali pa uporabljali ocenjevalno lestvico na različne načine. Leskošek idr. (2010) so ugotovili podobno, in

sicer da so v športni gimnastiki ob visoki zanesljivosti pogoste sistematične razlike med sodniki. Določeni sodniki namreč dosledno odstopajo pri ocenjevanju in kakovost sojenja se razlikuje glede na orodje, tekmovalno skupino in posameznega sodnika. Tudi Sato (2022) v analizi tekmovalnj v hip-hopu navaja, da so bile pristranskosti med sodniki značilne pri vseh ocenjevalnih kategorijah.

Vprašanje je, zakaj se opažene razlike v zanesljivosti pojavljajo tako med plesnimi zvrstmi kot med posameznimi sodniki. Eden od možnih razlogov za nekoliko nižjo zanesljivost v ST plesih je lahko večja kompleksnost pri ocenjevanju tehnične izvedbe. Sodniki pri ST plesih imajo morda bolj različne poglede na to, kateri vidiki tehnike so najpomembnejši (npr. drža v primerjavi s postavitvijo stopal ali linij telesa), zato se pojavljajo razlike v ocenjevanju. Nasprotno pa so v LA plesih tehnične pomanjkljivosti (npr. izguba ravnotežja, neuskkljenost v paru) morda očitnejše, kar vodi v višjo mero soglasja med sodniki. Na skladnost ocen lahko vplivajo tudi drugi subjektivni dejavniki. Študije kažejo, da lahko na sodnikovo percepcijo vplivajo celostni vizualni vtis, telesna podoba in mimika plesalcev (Sato, 2022), kar lahko pri nekaterih sodnikih povzroči nehoteno pristranskost.

Kljub dobri zasnovanosti sodniškega sistema ne moremo mimo dejstva, da je za razlike v ocenjevanju med sodniki krivo predvsem favoriziranje svojih plesnih parov. V športnem plesu je velika težava to, da so sodniki hkrati tudi trenerji, kar seveda privede do tega, da sodnik bolje oceni plesni par, katerega trener je. Čeprav je ta problematika žgoča že dolgo, sprememb ne načrtujejo. Poleg tega v športnem plesu ostajata dve močni plesni zvezi, in sicer World DanceSport federation in World Dance Council, ki se razlikujeta v poučevanju tehnike, načinu interpretacije in muzikalnosti. Zvezi imata različne poglede na razvoj športnega plesa, saj ena poudarja bolj »športno« plat, druga pa bolj »umetniško«, zato se stil plesanja med plesnimi pari različnih zvez razlikuje. Vpliv in moč posameznih sodnikov, ki zasedajo pomembne funkcije v zvezah, sta premočno izražena ter vplivata tudi na ocenjevanje

in pripadnost sodnikov eni ali drugi zvezi, to pa na končne rezultate, ki so pogosto nekorektni. Z doslednim upoštevanjem sodniških kriterijev bi bilo te razlike večinoma mogoče odpraviti, s čimer bi bilo tudi ocenjevanje bolj korektno in bi izražalo dejansko kakovost plesalcev.

Za prihodnji razvoj sodniškega sistema v ST in LA plesih so naše ugotovitve pomembna povratna informacija. Čeprav trenutni sistem že dosega razmeroma dobro zanesljivost, bi si bilo treba prizadevati za še večjo doslednost in objektivnost. Na podlagi naših ugotovitev in primerjav z drugimi športi priporočamo še natančnejšo opredelitev ocenjevalnih meril za TK in PV, posebej za standardne plese, kjer smo zaznali nekoliko slabšo zanesljivost. Poleg tega so pomembna redna usposabljanja sodnikov, pa tudi delavnice in analiza posnetkov plesnih nastopov, ki lahko sodnikom pomagajo, da uskladijo merila ocenjevanja. Nujni sta razmejitev vloge sodnika in trenerja, pri čemer posameznik ne sme opravljati obeh vlog, ter doslednost pri upoštevanju sodniških kriterijev ne glede na pripadnost eni od plesnih zvez. Smiselno bi bilo tudi uvesti mehanizme za zaznavanje in odpravo pristranskosti in izločitve skrajnih ocen ter zagotoviti večje število sodnikov na tekmovaljih. Navedeni predlogi bi lahko zmanjšali razlike v doslednosti in zanesljivosti ocenjevanja ter ga naredili še bolj pravičnega.

SKLEP

Študija je sicer pokazala večinoma dobro zanesljivost sodniškega ocenjevanja v vseh ST in LA plesih, a hkrati tudi nekatere pomanjkljivosti, in sicer nižjo zanesljivost TK in PS v ST plesih v primerjavi z LA plesi ter individualne razlike med sodniki. Te pomanjkljivosti so izziv za nadaljnji razvoj in izboljšanje sodniškega sistema ter dosledno in nepristransko ocenjevanje. Z izboljšavami za večjo objektivnost, kot so upoštevanje ocenjevalnih meril, redna usposabljanja sodnikov in sankcioniranje pristranskosti, lahko še izboljšamo zanesljivost ocenjevanja ter s tem pripomoremo k večjemu zaupanju plesalcev, trenerjev in javnosti v pravičnost sojenja v tekmovalnem plesu.

LITERATURA

- Atiković, A., Kalinski, S. D., Bijelić, S. in Vukadinović, N. A. (2011). Analysis results judging world championships in men's artistic gymnastics in the London 2009 year. *Sport. Log.* 7, 95–102. doi: 10.5550/sgia.110702.en.095A
- Boen, F., Van Hoya, K., Vanden Auweele, Y., Feys, J. in Smits, T. (2008). Open feedback in gymnastic judging causes conformity bias based on informational influencing. *Journal of sports sciences*, 26(6), 621–628.
- Bučar, M., Čuk, I., Pajek, J., Karacsony, I. in Leskošek, B. (2012). Reliability and validity of judging in women's artistic gymnastics at university games 2009. *Eur. J. Sport Sci.* 12, 207–215. doi: 10.1080/17461391.2010.551416
- Leskošek, B., Čuk, I., Karacsony, I., Pajek, J. in Bučar, M. (2010). Reliability and validity of judging in men's artistic gymnastics at the 2009 university games. *Sci. Gymnast J.* 2, 25–34.
- Lockwood, K. L., McCreary, D. R. in Liddell, E. (2005). Evaluation of success in competitive figure skating: an analysis of interjudge reliability. *Avante* 11, 1–9.
- Pajek, M. B., Čuk, I., Pajek, J., Kovač, M. in Leskošek, B. (2013). Is the quality of judging in women artistic gymnastics equivalent at major competitions of different levels? *J. Hum. Kinet.* 37, 173–181. doi: 10.2478/hu-kin-2013-0038
- Pajek, M. B., Čuk, I., Pajek, J., Kovač, M. in Leskošek, B. (2014). The judging of artistry components in female gymnastics: a cause for concern? *Sci. Gymnast J.* 6, 5–12.
- Plessner, H. in Schallies, E. (2005). Judging the cross on rings: A matter of achieving shape constancy. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 19(9), 1145–1156.
- Premelč, J. (2022). Vpliv števila sodnikov na zanesljivost sodniškega ocenjevanja v standardnih in latinskoameriških plesih. *Revija Šport*, 70.
- Premelč, J., Vučković, G., James, N. in Leskošek, B. (2019). Reliability of judging in DanceSport. *Frontiers in psychology*, 10, 1001.
- Sato, N. in Hopper, L. S. (2021). Judges' evaluation reliability changes between identifiable and anonymous performance of hip-hop dance movements. *PLoS ONE*, 16 (1).
- Singina, N. (2021). National preferences of referees in the first rounds of the dance discipline "Latin American Dances" of the WDSF World Championship. *Proceedings of the Russian State University of Physical Education, Sport, Youth and Tourism*, 1–8.
- WDSF (2024). Judging system. <https://old.worlddancesport.org/Rule/Competition/General/Judging-Systems>, pridobljeno 12. 9. 2025

dr. Jerneja Premelč, doc.
Univerza v Ljubljani, Fakulteta za šport
Jerneja.premelc@fsp.uni-lj.si