



# Let's get ready to rumble: Personality and the architecture of sport fandom

Dritjon Gruda<sup>a,b,\*</sup>, Aditya Simha<sup>c</sup>

<sup>a</sup> Universidade Católica Portuguesa, Católica Porto Business School, Research Centre in Management and Economics, Porto, Portugal

<sup>b</sup> Maynooth University, School of Business, Maynooth, Ireland

<sup>c</sup> University of Wisconsin–Whitewater, WI, USA

## ARTICLE INFO

### Keywords:

Personality  
Big Five  
Sport spectatorship  
Fandom  
Cultural omnivorousness  
Social media language

## ABSTRACT

Who watches sport, and what does personality reveal about how fandom is expressed? Using Facebook status updates and self-reported Big Five scores of 6246 users across 17 sports, we show that personality leaves a small but consistent imprint on the expression of sport fandom while barely predicting which sport a person watches. Personality maps onto sport attributes more than sport identities: less agreeable users favored high-contact, high-risk sports and extraverts favored larger-team sports. The breadth of a person's viewing was associated with Extraversion rather than Openness, contrary to the cultural omnivore hypothesis. Fans' own words, scored for sentiment and for motives with a validated language-model classifier, show that personality predicts the intensity and motives of expressed fandom (drama, escape, social connection). These associations hold when posting volume and inferred age and gender are controlled, and we found no reliable evidence that inferred gender moderates them. Personality shapes the expression of engagement more than taste in sports.

Billions of people watch sports. Some 1.4 billion people watched the 2022 World Cup final alone, and the 2026 tournament already broke broadcast records (FIFA, 2024, 2026). Psychology is not silent about these viewers. Research has mapped fan motives (Trail & James, 2001; Wann, 1995), team identification (Wann & Branscombe, 1993), and the personality of spectators (Appelbaum et al., 2012; Kim et al., 2017; Ko et al., 2017). But this literature rests largely on modest self-report surveys, typically within a single sport or fan community. What it cannot yet say is how personality relates to spectatorship at scale, across many sports at once, and in fans' own spontaneous words rather than questionnaire responses. That is the gap: not an absence of research, but an absence of scale, cross-sport comparison, and spontaneous language.

In this paper, we examine two questions. The first is direct: do people higher on particular Big Five traits prefer to watch particular sports? The second is broader. Beyond the choice of a sport, does personality shape the wider architecture of fandom? By *architecture of fandom* we mean four measurable components of how fandom is expressed. These are the attributes of the sports a person mentions watching, the breadth of sports followed, the valence and intensity of sport talk, and the motives that talk expresses. We follow the fan literature in treating fandom itself as passionate and often collective engagement with sport, comprising identification with teams (Wann & Branscombe, 1993), emotional investment (Cialdini et al., 1976), and distinct consumption motives

(Wann, 1995). One caution frames everything that follows. Our unit of evidence is *expressed* fandom (sport-related language produced spontaneously on Facebook), not verified viewing behavior, a stable measured preference, or a validated fandom scale.

Because broad dispositions tend to predict broad tendencies rather than narrow choices (Funder & Ozer, 2019), the second question may prove the more informative. We make four contributions. First, we relate the personality profiles of a sport's audience to the attributes of the sport itself. Second, we ask whether personality predicts the breadth of a person's sport viewing, connecting the analysis to the omnivore thesis in the sociology of taste (Peterson & Kern, 1996; Widdop et al., 2016). Third, because our data are people's own words, we ask with what intensity and identification people talk about sport and which spectator motives their language expresses (Trail & James, 2001; Wann, 1995). Fourth, we derive a personality-based typology of sports from their audiences' trait profiles.

## 1. Personality, athletes, and fans

The Big Five offer a parsimonious framework for individual differences, and decades of research connect them to sport. Allen et al. (2013) reviewed seven decades of research and concluded that personality differentiates athletes from non-athletes, with team and high-risk

\* Corresponding author at: Católica Porto Business School, Porto, Portugal.

E-mail address: [jgruda@ucp.pt](mailto:jgruda@ucp.pt) (D. Gruda).

<https://doi.org/10.1016/j.paid.2026.114030>

Received 23 June 2026; Received in revised form 27 July 2026; Accepted 2 August 2026

Available online 4 August 2026

0191-8869/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

athletes tending to score higher on Extraversion and lower on Conscientiousness than individual-sport athletes. Language analytics extend this line to psychological states: pre-game anxiety expressed in athletes' tweets predicts NBA performance (Gruda & Ojo, 2023).

Research on fans is thinner, but real. Russell and Goldstein (1995) reported personality differences between football fans and non-fans; Appelbaum et al. (2012) linked Extraversion, specifically excitement-seeking and gregariousness, to spectating; and Kim et al. (2017) found distinct personality clusters among users of team social networking sites. This work establishes that personality matters in sport but leaves the mechanism open and largely treats fandom as a single, undifferentiated act.

## 2. From which sport to how and why: the style and motives of fandom

Sport-consumer research has long held that the same behavior, watching a game, can serve different psychological ends. The Sport Fan Motivation Scale (Wann, 1995) and the Motivation Scale for Sport Consumption (Trail & James, 2001) distinguish motives including drama and eustress, escape, social interaction, aesthetics, and vicarious achievement. A parallel tradition documents how fans manage identity through their teams: basking in reflected glory after wins (Cialdini et al., 1976), with team identification varying widely across individuals (Wann & Branscombe, 1993). Personality predicts both team identification and its underlying motives (Brown-Devlin & Devlin, 2020), and a hierarchical strand treats personality as a distal antecedent of consumption that works through psychological needs such as arousal and affiliation (Ko et al., 2017). Personality may thus shape the experience and function of fandom even when it weakly predicts the object of fandom.

Peterson and Kern (1996) showed that high-status cultural consumers have shifted from snobbish exclusion toward omnivorous appreciation of many forms, and Openness is the trait most consistently tied to broad cultural taste (Rentfrow & Gosling, 2003). The omnivore thesis has been extended to sport: watching a wide range of sports is associated with social capital and networks rather than cultural capital alone (Widdop et al., 2016). Does sport omnivory rest on the openness of the cultural omnivore or on the sociability of the connected fan?

These strands point to a single organizing idea. If personality shapes spectatorship, it should do so not sport by sport but through a few broad dispositional channels. We propose two. The first is *arousal regulation*: the disposition to approach the high-arousal, combative, crowd-charged spectacle that team and contact sports stage. Excitement-seeking (a facet of Extraversion), tolerance for aggression and dominance (the low pole of Agreeableness), and an approach orientation toward charged social settings (the low pole of Neuroticism) should load here. The second is *social affiliation*: the disposition to use sport as a vehicle for connection. Sociability should express itself less in the choice of a sport than in the breadth of sports followed and in motives of social connection and identification. We treat these channels as an organizing interpretation, not a structure estimated independently of the patterns they summarize; arousal and affiliation are two of the needs Ko et al. (2017) placed between personality and sport consumption.

These channels, with the literatures above, support four directional hypotheses. Extraversion entails sociability, gregariousness, and excitement-seeking, and both team athletes and spectators have shown elevated Extraversion (Allen et al., 2013; Appelbaum et al., 2012):

**H1.** Extraversion is positively associated with a preference for team over individual sports.

The omnivore thesis holds that wide-ranging consumption marks cultural status (Peterson & Kern, 1996), and Openness is the trait most tied to breadth of taste (Rentfrow & Gosling, 2003):

**H2.** Openness is positively associated with the breadth of sports a

person watches.

Extraversion is defined in part by a need for social interaction, and social affiliation is a recognized spectator motive (Trail & James, 2001; Wann, 1995):

**H3.** Extraversion is positively associated with watching for social motives.

Excitement-seeking is a core facet of Extraversion, and drama, or eustress, is among the most cited fan motives (Wann, 1995), with excitement-seeking tied specifically to spectating (Appelbaum et al., 2012):

**H4.** Extraversion is positively associated with watching for drama/eustress motives.

## 3. The present research

Where Ko et al. (2017) surveyed 471 consumers and modeled how personality raises consumption involvement through latent needs, we use the spontaneous language of 6246 users, from which personality traits can be reliably recovered (Karanatsiou et al., 2022), to ask where, across expressed fandom, personality leaves its mark. H1–H4 are directional tests stated in advance, though the study is observational and not preregistered, so even these are evidence rather than confirmation. All remaining analyses (the sport-attribute analysis, the sentiment and expression models, motive effects beyond H3 and H4, and the sport typology) are exploratory and labeled as such in the Results.

## 4. Method

### 4.1. Sample and procedure

Data are from the myPersonality project, a corpus of Facebook users who completed psychometric questionnaires and consented to research use of their profile data and self-reported Big Five scores (cf. Kosinski et al., 2013). From an initial random sample of 72,748 participants, we identified status updates mentioning any of 17 popular sports and retained the 6246 users who mentioned at least one sport in a spectatorship context and had complete Big Five scores. The corpus includes no demographic fields for these users; we therefore estimated age and gender from language (below).

### 4.2. Measures

**Sport preference.** For each of 17 sports (athletics, badminton, baseball, basketball, boxing, cricket, football, golf, gymnastics, hockey, marathon, rugby, soccer, swimming, tennis, volleyball, wrestling) we coded a binary indicator of whether a user mentioned that sport in a spectatorship context. A mention indicates expressed interest (the salience of a sport in a person's own words), not verified viewing or a stable measured preference; we use “watch” as shorthand for this expressed spectatorship. Breadth (sport omnivory) is the count of distinct sports a user mentioned. Team sports were distinguished from individual sports by team size (Supplementary Materials S7).

**Personality.** The Big Five were measured by the self-reported 20-item International Personality Item Pool scale (Goldberg et al., 2006) on a 1–5 metric: Openness ( $M = 3.79$ ,  $SD = 0.68$ ), Conscientiousness ( $M = 3.49$ ,  $SD = 0.73$ ), Extraversion ( $M = 3.59$ ,  $SD = 0.80$ ), Agreeableness ( $M = 3.59$ ,  $SD = 0.70$ ), and Neuroticism ( $M = 2.68$ ,  $SD = 0.79$ ). Traits were standardized for analysis. The archival data provide trait scores rather than item responses, so internal-consistency reliabilities could not be recomputed here. Brief scales carry more measurement error than full-length instruments, and random error of that kind attenuates observed associations. Because reliability could not be estimated in these data, however, the size and even the direction of any resulting bias in the effects reported below are unknown.

**Spectatorship sentiment (valence).** For each sport mention we located the status line(s) naming the sport and scored their sentiment with VADER, a lexicon-and-rule model validated for social-media text (Hutto & Gilbert, 2014). A compound score at or above 0.05 was coded positive, at or below  $-0.05$  negative, otherwise neutral. A sport-context filter removed non-sport uses of sport keywords (e.g., movie marathons, “boxing day,” cricket the insect), cutting the off-topic rate from 10.8% to 0.8% at perfect precision and 93% recall against hand coding.

**Fan expression.** From each user's sport lines we computed emotional intensity (mean absolute VADER compound), exclamation-mark and ALL-CAPS rates, and first-person-plural pronoun use (“we/us/our”) as a marker of team identification (Cialdini et al., 1976; Wann & Branscombe, 1993). Because first-person-plural usage also arises in ordinary conversation, we treat it as a coarse proxy rather than a validated measure.

**Fan motives.** Guided by the Motivation Scale for Sport Consumption (Trail & James, 2001), we coded five motives: social, drama/eustress, escape, aesthetic, and vicarious achievement. We wrote an explicit codebook, reproduced with its decision rules and anchoring examples in Supplementary Materials S6, and applied it at scale with a five-round, majority-vote large-language-model classifier (Claude Haiku 4.5) over all 13,061 on-topic spectatorship posts. Coding is multi-label: a post may carry several motives or none, and posts meeting no decision rule were left uncoded rather than assigned a default. The 154-post validation sample was drawn from the on-topic lines by stratified sampling designed to yield enough positives for every motive, so it over-represents coded posts by design. The classifier was highly self-consistent across rounds (Fleiss' kappa 0.91 to 0.97) and, against a two-coder consensus gold standard built from the same codebook (Supplementary Materials S6), reached substantial agreement for vicarious (kappa = 0.87), social (0.79), and aesthetic (0.79) and moderate agreement for drama/eustress (0.62) and escape (0.52), outperforming a keyword-dictionary baseline (kappa 0.22 to 0.52) on every motive except escape, where the two are equal. Agreement is weakest for escape, the most inferential category; we read escape and the rare aesthetic motive (2.6% of users) with corresponding caution.

**Inferred demographics.** Because the corpus lacks demographic fields, we estimated each user's age and gender from their full status text using the predictive lexica of Sap et al. (2014), which predict held-out age at  $r = 0.83$  and gender at 91.9% accuracy in social-media samples. The resulting estimates (inferred mean age 23.4 years; 57% inferred female) match the known profile of the myPersonality corpus. These are language-derived proxies with error. Models including them are sensitivity checks against demographic composition, not a substitute for adjustment on measured demographics.

**Sport attributes.** We coded each of the 17 sports on eight attributes: physical contact, combat, injury risk, aesthetic/judged character, pace, team size, mainstream popularity, and indoor/outdoor. Ratings were made by the author team by consensus against written rules, with disagreements resolved by discussion. Because the procedure produced no independent ratings, no agreement coefficient can be computed from it, and the complete coding is reported in Supplementary Materials S7. Resting on consensus ratings of 17 cases, some of record (team size, indoor or outdoor) and some of judgment (injury risk, pace, aesthetic character), the analysis is exploratory and theory-generating.

#### 4.3. Analytic strategy

We estimated point-biserial trait-by-sport correlations and a team-preference logistic regression, then evaluated robustness three ways: Benjamini-Hochberg false-discovery-rate control across the 85 trait-by-sport tests (Benjamini & Hochberg, 1995); per-sport multivariate logistic regressions entering all five standardized traits; and three operationalizations of the team/individual contrast. Breadth was modeled with Poisson regression controlling for posting volume. Sentiment was modeled with sport fixed effects and user-clustered standard errors.

Sports were clustered on audience trait profiles (Ward's method), and co-viewing was related to profile similarity controlling for popularity. The same false-discovery-rate correction was applied within each text-derived family. All headline models were re-estimated adding inferred age and gender as covariates, and trait  $\times$  gender interactions were tested (Supplementary Materials S3 and S5). Effect sizes are interpreted against the benchmarks of Funder and Ozer (2019).

## 5. Results

Table 1 reports descriptive statistics and intercorrelations. In the team-preference logistic regression, Extraversion is positively associated with watching team sports ( $b = 0.16$  per SD,  $p < .001$ ), while Openness ( $b = -0.11$ ), Agreeableness ( $b = -0.08$ ), and Neuroticism ( $b = -0.11$ ) are negative, with Conscientiousness nonsignificant ( $N = 6246$ ). Because a naive coding equates team viewing with mentioning any team sport, we estimated the contrast three ways (Fig. 2). Consistent with H1, the pattern is stable whether team preference is defined as any team sport, as a pure-type contrast dropping mixed viewers, or as preference expressed with positive sentiment. In predicted probabilities, a two-SD span of Extraversion separates 63.5% from 70.4% mentioning a team sport, against a base rate of 67.0%.

Predicting the number of distinct sports a person watches, controlling for posting volume, Extraversion is associated with a modestly broader sport portfolio (IRR = 1.03 per SD,  $p = .007$ ), while Openness is associated with a slightly narrower one (IRR = 0.97 per SD,  $p = .010$ ; Fig. 4). Contrary to H2, Openness did not predict broader viewing. In expected counts, a two-SD span of Extraversion separates 1.39 from 1.48 distinct sports, against a mean of 1.44. The effects are small, but their direction runs counter to the cultural-omnivore prediction and fits the social-capital account of Widdop et al. (2016) better than a cultural-capital one, neither of which we measured.

The motive results, from the validated language-model classifier (Fig. 5), support H3 and H4: Extraversion predicts both the social motive (OR = 1.18 per SD) and the drama/eustress motive (OR = 1.13), and both effects survive false-discovery-rate correction (Supplementary Materials). Across the same two-SD span, 30.1% versus 37.5% express the social motive and 29.3% versus 34.7% the drama motive: reliable, but modest in absolute terms.

#### 5.1. Exploratory analyses

Of the 85 trait-by-sport correlations, 24 reach  $p < .05$  and only 14 survive false-discovery-rate control. Per-sport multivariate models give a clearer summary (Fig. 1). Entering all five traits at once, only 12 of 85 coefficients survive correction, and they cohere: Extraversion raises the odds of watching the popular team sports (football, soccer, basketball) and golf, and lowers them for marathon and tennis; Openness is higher for marathon and lower for basketball and gymnastics; Neuroticism is higher for swimming.

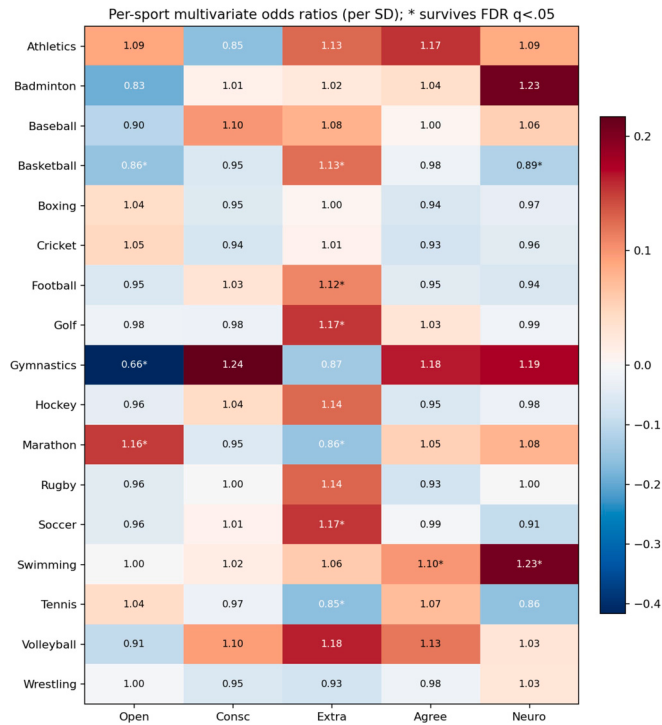
Relating each sport's audience trait profile to its attributes organizes this scatter (Fig. 3). Less agreeable audiences favor high-contact ( $\rho = -0.56$ ) and high-risk ( $\rho = -0.60$ ) sports; Extraversion tracks larger-team sports ( $\rho = 0.56$ ); Neuroticism tracks niche ( $\rho = -0.63$  with popularity), small-team ( $\rho = -0.52$ ), and aesthetic ( $\rho = 0.51$ ) sports. With only 17 consensus-rated sports (Supplementary Materials S7), these associations are exploratory and cannot establish contact, risk, team size, or popularity as mechanisms; we advance them as hypotheses for future work, though they suggest personality responds to properties of a sport rather than to its name.

A mention is not a preference. Across cleaned spectatorship lines, 59.9% carry positive sentiment, 22.6% are neutral, and 17.5% are negative; presence/absence coding silently merges enthusiasm with complaint. Sentiment also varies with personality. Controlling for which sport is discussed and clustering by user, Agreeableness ( $b = 0.025$ ), Conscientiousness (0.022), Extraversion (0.021), and Neuroticism

**Table 1**  
Descriptive statistics and intercorrelations among study variables ( $N = 6246$ ).

| Variable             | <i>M</i> | <i>SD</i> | 1     | 2     | 3     | 4     | 5     | 6    | 7    |
|----------------------|----------|-----------|-------|-------|-------|-------|-------|------|------|
| 1. Openness          | 3.79     | 0.68      |       |       |       |       |       |      |      |
| 2. Conscientiousness | 3.49     | 0.73      | 0.01  |       |       |       |       |      |      |
| 3. Extraversion      | 3.59     | 0.80      | 0.11  | 0.16  |       |       |       |      |      |
| 4. Agreeableness     | 3.59     | 0.70      | 0.03  | 0.18  | 0.13  |       |       |      |      |
| 5. Neuroticism       | 2.68     | 0.79      | −0.04 | −0.28 | −0.33 | −0.32 |       |      |      |
| 6. Breadth           | 1.44     | 0.79      | −0.04 | 0.02  | 0.09  | 0.02  | −0.03 |      |      |
| 7. Team preference   | 0.67     | 0.47      | −0.04 | 0.04  | 0.08  | −0.01 | −0.07 | 0.30 |      |
| 8. Sentiment         | 0.26     | 0.42      | −0.03 | 0.06  | 0.05  | 0.06  | −0.01 | 0.01 | 0.04 |

Note. Team preference and Sentiment are a binary team-sport indicator and the mean valence of a user's sport mentions, respectively. Correlations  $|r| \geq 0.03$  are significant at  $p < .05$ .



**Fig. 1.** Per-sport multivariate odds ratios per SD of each trait. Note. Each sport is regressed on all five standardized traits. Asterisks denote coefficients surviving false-discovery-rate control.

(0.017) each predict more positive sport talk, while Openness predicts slightly less positive talk (−0.011). Team sports are discussed more warmly than individual sports (mean compound 0.27 vs 0.24,  $p = .003$ ). Personality is also visible in how people talk about sport (Table S2). Extraversion predicts greater emotional intensity ( $b = 0.013$  per  $SD$ ,  $p < .001$ ), more exclamation marks ( $b = 0.153$ ,  $p < .001$ ), more ALL-CAPS ( $b = 0.117$ ,  $p = .001$ ), and more collective “we” language ( $b = 0.074$ ,  $p = .007$ ); Conscientiousness, Agreeableness, and Neuroticism also predict more “we” language, and Openness predicts fewer exclamations ( $b = −0.087$ ,  $p = .007$ ). We read these as convergent indicators of the intensity and collective voice of expressed fandom, not as a validated fandom scale. Among the remaining motives, escape tracks Extraversion, Agreeableness, and Neuroticism (ORs near 1.10); the social motive also rises with Conscientiousness, Neuroticism, and Agreeableness and falls with Openness (0.92); vicarious achievement is most strongly predicted by Conscientiousness (1.20); the rare aesthetic motive shows no reliable trait association.

Clustering sports by their audiences' Big Five profiles yields a two-family structure: mainstream team and ball sports (football, soccer, basketball, baseball, hockey, rugby, volleyball, with golf and tennis) versus endurance, combat, and niche sports (boxing, wrestling,

marathon, swimming, badminton, cricket, athletics), with gymnastics an outlier. The clustering crosses the team/individual line. Psychologically similar sports are not, however, especially co-viewed once sport popularity is controlled ( $b = −0.001$ ,  $p = .67$ ): a bounded negative result.

**Robustness.** Three checks bound these claims. First, the patterns are not an artifact of posting volume. Extraverts post somewhat more ( $r = 0.10$ ), but the Extraversion effects on the social and drama motives hold within strata of posting volume and under a quadratic volume control. Second, the even spread of personality's imprint is not a reliability illusion: measured at the same level of aggregation, the five traits explain a comparable share of variance in each (about 1%), and the most reliably measured outcome, team orientation, is the most strongly predicted. Third, the results are not explained by inferred demographics. Adding estimated age and gender leaves Extraversion's associations with team preference, breadth, expression, and the social and drama motives essentially unchanged, while the Neuroticism and Agreeableness associations with the social motive and sentiment attenuate to non-significance and the Conscientiousness sentiment effect halves (Tables S3–S4). Some affect-related effects thus partly reflect demographic composition; the Extraversion architecture does not. Nor did we find reliable evidence of moderation by inferred gender: none of 55 trait  $\times$  inferred-gender interactions across the headline models survives false-discovery-rate correction (Supplementary S5). Because gender was inferred from language rather than measured, misclassification attenuates interaction estimates, so this null is weak evidence of homogeneity. Gender differences in expressed fandom appear here as differences of level rather than of trait mechanism. Residual confounding remains possible.

6. Discussion

Three patterns organize the results. First, and most tentatively, the pull of a sport appears to run through its attributes. Less agreeable viewers are drawn to harder, more combative sports, complicating the cooperation-and-camaraderie story usually attached to team sport. Resting on consensus ratings of 17 sports, this is a hypothesis about sport attributes, not an established effect. Second, to the extent that breadth of sport viewing is dispositional at all, it tracks Extraversion rather than Openness, which fits the social-capital account of Widdop et al. (2016) better than a cultural-capital one, though neither form of capital was measured here. Third, personality's imprint is spread across the architecture of expressed fandom rather than concentrated in one place. The kind of sport, the breadth of viewing, and the intensity and motives of engagement are each predicted to a similar, modest degree, while the specific sport is barely predicted at all.

Read together, these patterns are consistent with the two dispositional channels we proposed, though neither channel was measured, and both remain interpretations rather than demonstrated processes. If arousal regulation organizes what people watch, extraverted, less agreeable, and less neurotic viewers should gravitate toward the combative, high-arousal spectacle of team and contact sports and should

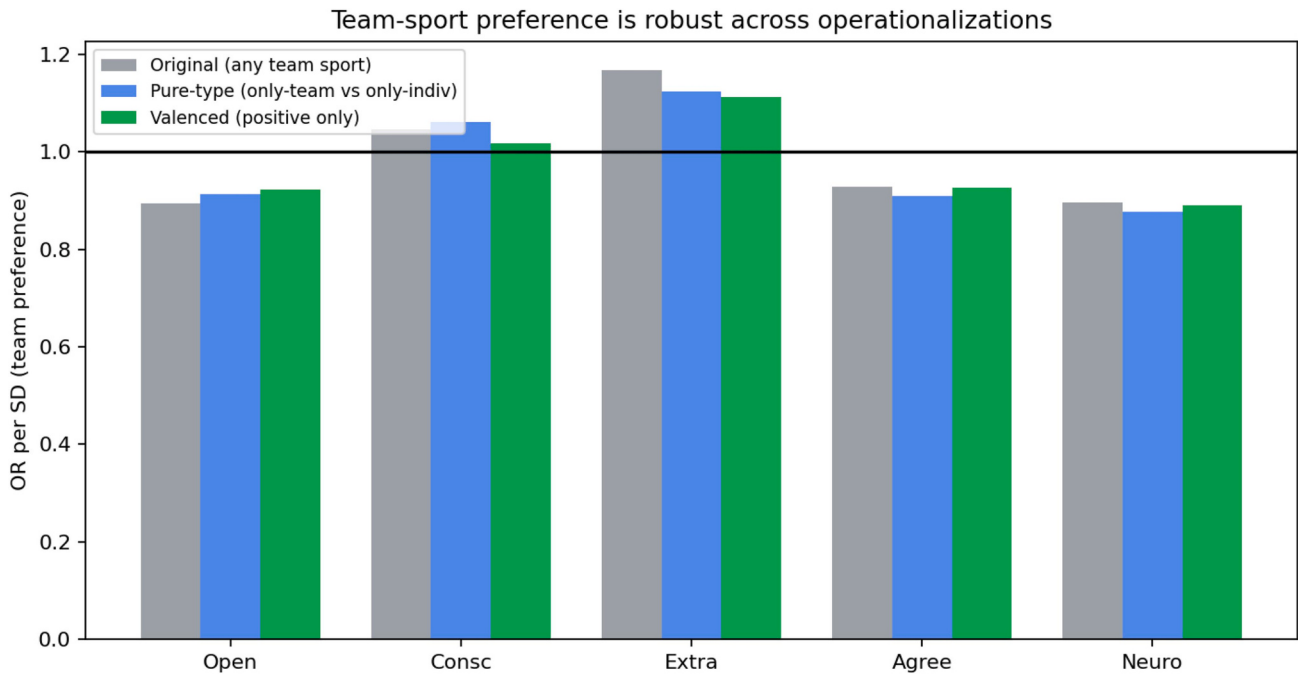


Fig. 2. Team-preference odds ratios per SD across three operationalizations.

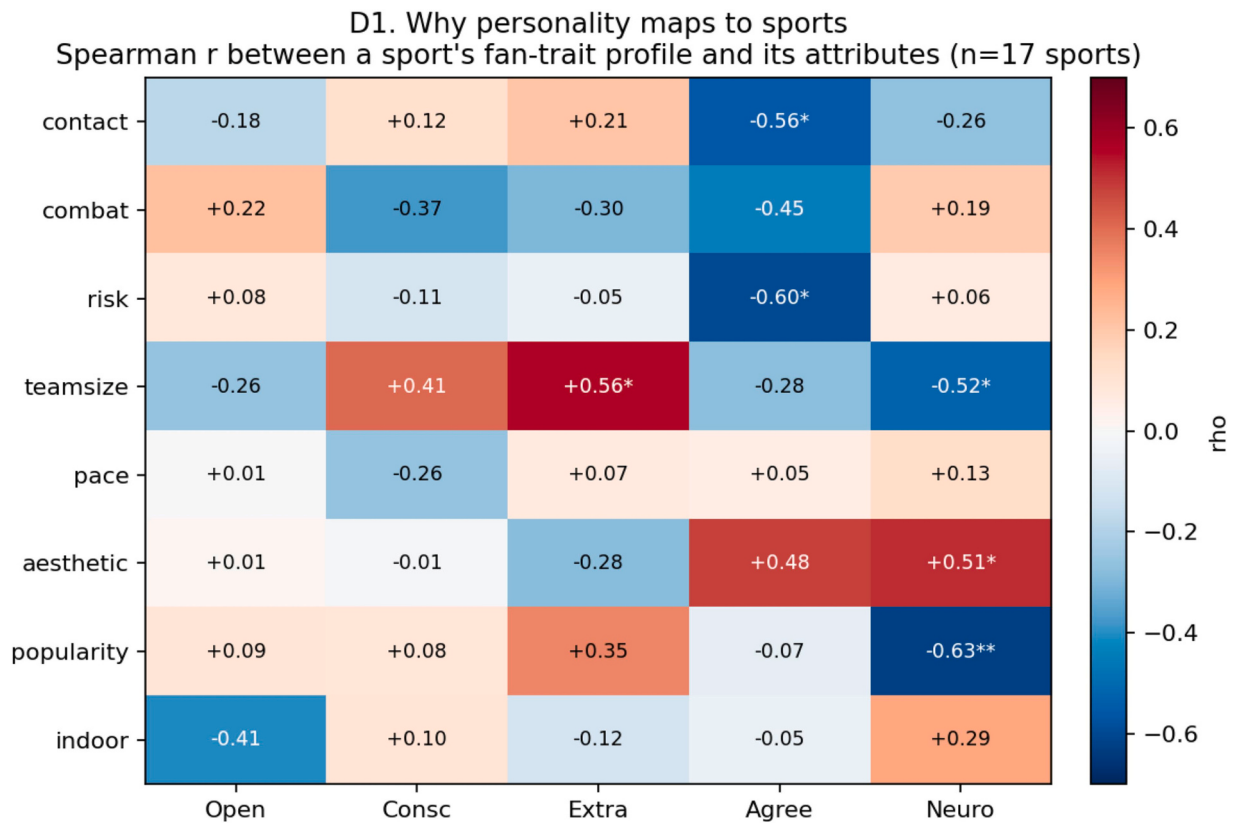


Fig. 3. Spearman correlations between sport attributes and sports' fan-trait profiles. Note.  $n = 17$  sports; exploratory.

report watching for drama, which is what we observe. If social affiliation organizes how broadly and why they watch, sociability should widen the range of sports followed and accompany motives of connection and identification, which it does. No index of arousal or affiliation was assessed, so both readings remain candidate explanations. Openness sits

outside both channels: more open users showed slightly narrower viewing, less positive sport talk, and lower social and vicarious motives, a consistent, small pattern of lower engagement with mainstream spectacle. These are small associations by conventional benchmarks (Funder & Ozer, 2019), and we frame them accordingly. Personality

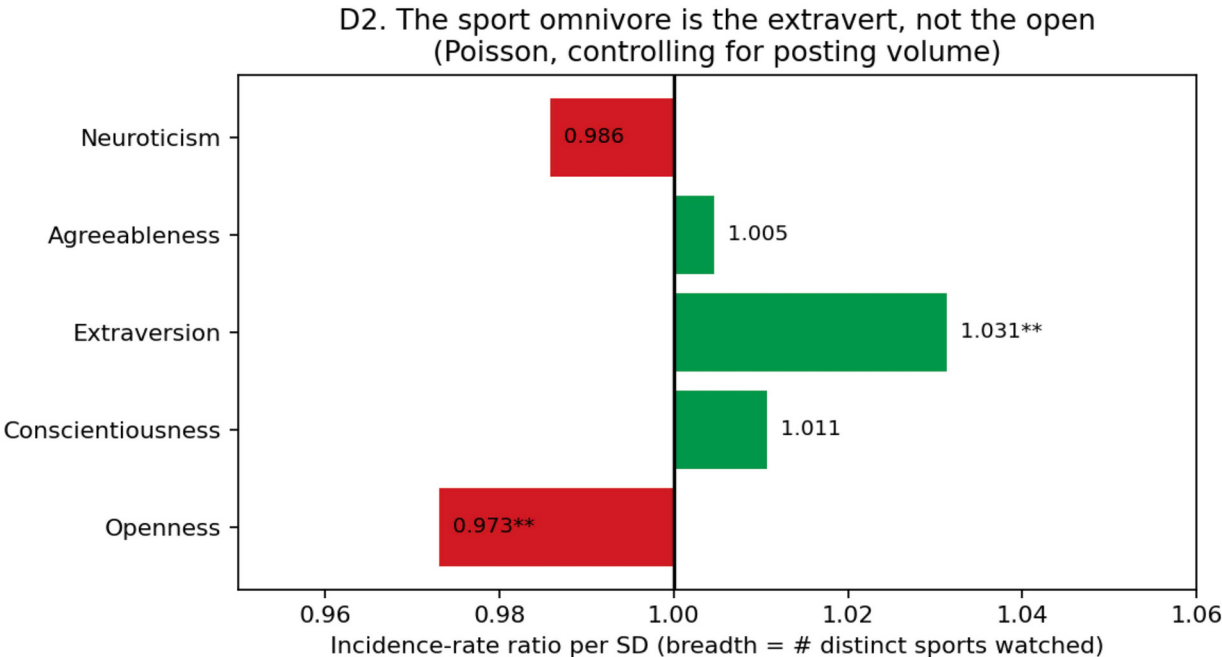


Fig. 4. Breadth of sports watched: incidence-rate ratio per SD.  
Note. Poisson regression with posting volume controlled.

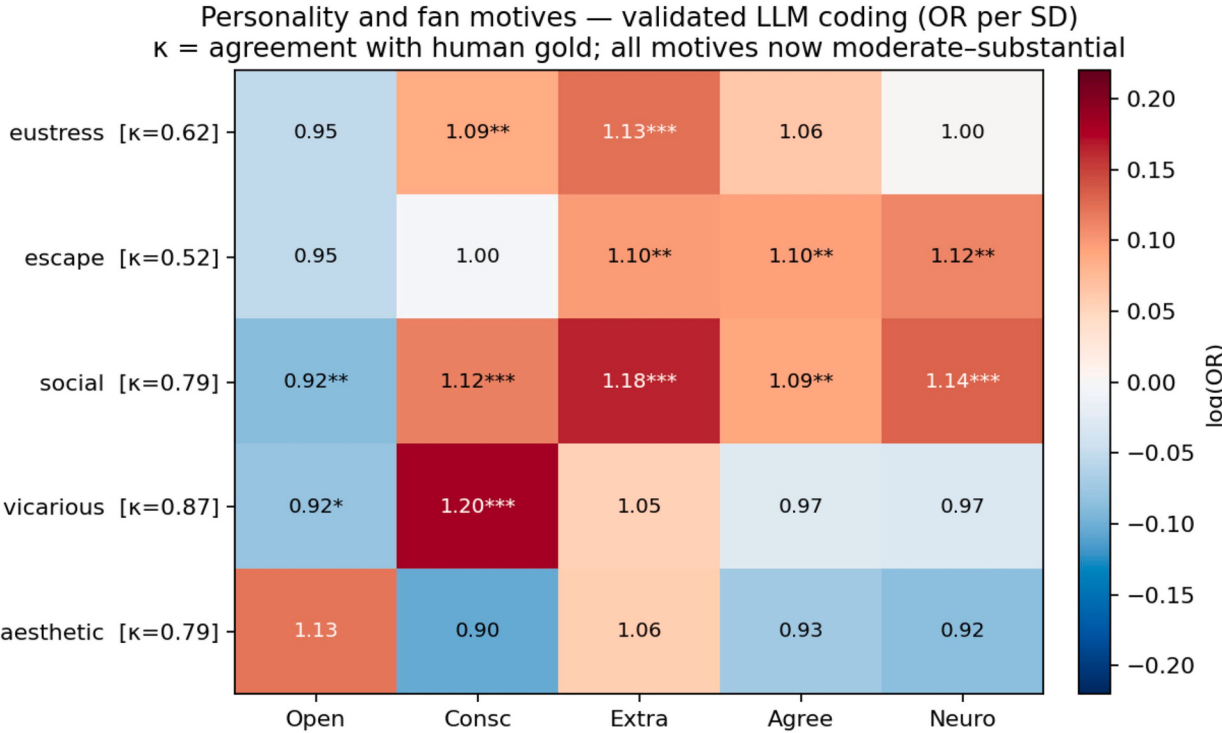


Fig. 5. Personality and fan motives (odds ratios per SD).  
Note. Motives are from a five-round, majority-vote large-language-model classifier validated against human coding; bracketed values are Cohen's kappa versus the human gold standard. Asterisks denote significance.

shades the expression of fandom; it does not determine it.  
If personality predicts the shape of engagement (attributes, breadth, motives) rather than the specific object, then designs fixed on a single sport will understate personality's role, while designs that aggregate across the field will recover it, echoing the aggregation principle (Funder & Ozer, 2019). Practically, and speculatively given the small

effects, engagement strategies keyed to the shape of fandom (breadth, intensity, motives) rather than to a single sport may be worth testing.

6.1. Limitations and future directions

Several limitations bound these conclusions. The data are cross-

sectional and historical, drawn from a now-retired corpus, so neither causal nor contemporary claims are warranted; our contribution is a proof-of-concept that self-reported personality, spontaneous language, and sport mentions can be combined to show where personality does and does not organize expressed spectatorship. The corpus contains no measured demographics. We addressed this by inferring age and gender from language with validated lexica and re-estimating all headline models, which left the Extraversion results unchanged while attenuating some affect-related effects, but inferred covariates are error-prone substitutes, and nationality, culture, and social class remain unmeasured; all plausibly shape both sport interest and posting behavior. Sport preference is inferred from spontaneous mention, which captures salience rather than exhaustive viewing, and posting is non-random, so the expression and sentiment measures describe an expressive subsample. Personality was assessed with a brief 20-item IPIP scale whose reliabilities could not be recomputed. Random error attenuates associations, but without reliability estimates we cannot say by how much, and differential reliability across traits could also distort the relative sizes we compare. The text-derived measures carry known error. VADER sentiment is moderately valid (kappa approximately 0.46) and best read in aggregate, and the motive gold standard was reached by consensus between two coders reviewing the same posts rather than by independent double coding, so no human inter-coder reliability statistic is available.

Finally, the Big Five were the only personality measures available in this corpus, and broadband traits are not necessarily the individual differences most proximal to fandom. Narrower constructs (competitiveness, status seeking, the tendency to bask in reflected glory, tribalism, sensation seeking) may correlate more strongly with fandom's facets, and our arousal-regulation and social-affiliation channels are naturally operationalized by exactly such mid-level traits. Pairing them with validated fandom instruments in contemporary data is the clearest next step.

## 7. Conclusion

Personality shows small but consistent associations with how sport fandom is expressed online (the attributes of the sports people mention, the breadth of what they follow, and the intensity and motives of their sport talk), but it does not strongly predict which specific sport a person watches. Asking which sport people of a given personality watch yields thin answers; asking where, across expressed fandom, personality leaves its mark yields richer ones. The personality of fandom is less about the sport than about the fan.

## CRedit authorship contribution statement

**Dritjon Gruda:** Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Aditya Simha:** Writing – review & editing, Investigation, Conceptualization, Methodology.

## Ethics

The myPersonality project obtained informed consent for research use of participants' data, which were anonymized. This secondary analysis of fully anonymized archival data did not require additional ethical approval.

## Declaration of Generative AI and AI-assisted technologies in the writing process

A large language model (Claude Haiku 4.5) was used as a measurement instrument to code fan motives, as described in the [Method](#) section, under an explicit human-written codebook with reliability reported against human coding. All study design, analysis, interpretation, and final text are the authors' own.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Declaration of competing interest

The authors declare no competing interests.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.paid.2026.114030>.

## Data availability

The myPersonality data are not publicly available because the project is closed to new research access. De-identified derived data sufficient to reproduce the reported analyses, together with all analysis code (a Python pipeline and a Stata do-file), are available from the authors upon reasonable request.

## References

- Allen, M. S., Greenlees, I., & Jones, M. (2013). Personality in sport: A comprehensive review. *International Review of Sport and Exercise Psychology*, 6(1), 184–208.
- Appelbaum, L. G., Cain, M. S., Darling, E. F., Stanton, S. J., Nguyen, M. T., & Mitroff, S. R. (2012). What is the identity of a sports spectator? *Personality and Individual Differences*, 52(3), 422–427.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, 57(1), 289–300.
- Brown-Devlin, N., & Devlin, M. B. (2020). Winning with personality: Underscoring antecedents for college students' motives for team identification. *Communication & Sport*, 8(3), 364–388.
- Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S., & Sloan, L. R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology*, 34(3), 366–375.
- FIFA. (2024, November 29). Reports detail record global audience figures and positive sustainability outcomes from FIFA World Cup Qatar 2022. <https://inside.fifa.com/or- ganisation/news/reports-detail-record-global-audience-figures-and-positive-sustainability-outcomes-fifa-world-cup-qatar>.
- FIFA. (2026, June 17). FIFA World Cup breaks broadcast records across the board [Press release]. <https://inside.fifa.com/media-releases/world-cup-breaks-broadcast-record-s-across-board>.
- Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, 2(2), 156–168.
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1), 84–96.
- Gruda, D., & Ojo, A. (2023). All I do is win, win, win no matter what? Pre-game anxiety and experience predict athletic performance in the NBA. *Current Research in Behavioral Sciences*, 5, Article 100120.
- Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *8(1). Proceedings of the International AAAI Conference on Web and Social Media* (pp. 216–225).
- Karanatsiou, D., Sermpetzis, P., Gruda, D., Kafetsios, K., Dimitriadis, I., & Vakali, A. (2022). My tweets bring all the traits to the yard: Predicting personality and relational traits in online social networks. *ACM Transactions on the Web*, 16(2), 1–26.
- Kim, Y., Kim, S., & Kim, Y. M. (2017). Big-five personality and motivations associated with sport team social networking site usage: A cluster analysis approach. *Journal of Global Sport Management*, 2(4), 250–274.
- Ko, Y. J., Chang, Y., Jang, W., Sagas, M., & Spengler, J. O. (2017). A hierarchical approach for predicting sport consumption behavior: A personality and needs perspective. *Journal of Sport Management*, 31(3), 213–228.
- Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15), 5802–5805.
- Peterson, R. A., & Kern, R. M. (1996). Changing highbrow taste: From snob to omnivore. *American Sociological Review*, 61(5), 900–907.
- Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84(6), 1236–1256.
- Russell, G. W., & Goldstein, J. H. (1995). Personality differences between Dutch football fans and nonfans. *Social Behavior and Personality*, 23(2), 199–204.
- Sap, M., Park, G., Eichstaedt, J., Kern, M., Stillwell, D., Kosinski, M., ... Schwartz, H. A. (2014). Developing age and gender predictive lexica over social media. In

- Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1146–1151).
- Trail, G. T., & James, J. D. (2001). The motivation scale for sport consumption: Assessment of the scale's psychometric properties. *Journal of Sport Behavior*, 24(1), 108–127.
- Wann, D. L. (1995). Preliminary validation of the sport fan motivation scale. *Journal of Sport and Social Issues*, 19(4), 377–396.
- Wann, D. L., & Branscombe, N. R. (1993). Sports fans: Measuring degree of identification with their team. *International Journal of Sport Psychology*, 24(1), 1–17.
- Widdop, P., Cutts, D., & Jarvie, G. (2016). Omnivorousness in sport: The importance of social capital and networks. *International Review for the Sociology of Sport*, 51(5), 596–616.