Perceived quality of group activities fitness instructors: a comparative study across stakeholders, demographics, and club contexts

Calidad percibida de las actividades grupales de los instructores de fitness

Autores

Daniela Almeida [1]
Vera Simões [2, 3]
Susana Franco [2, 3]
José Aleixo [1]
Fernando Martins [1, 3, 4, 5]
Francisco Campos [1, 3]

[1] Polytechnic University of Coimbra (Portugal)
[2] Santarém Polytechnic University (Portugal)
[3] SPRINT - Sport Physical activity and health Research & INnovation cenTer (Portugal)
[4] Instituto de Telecomunicações, Delegação da Covilhã (Portugal)
[5] inED - Centre for Research & Innovation in Education (Portugal)

Received: 17-07-25
Accepted: 19-12-25

Autor de correspondência:
Francisco Campos
francicampos@esec.pt

How to cite in APA

Almeida, D., Simões, V., Franco, S., Aleixo, J., Campos, F., & Martins, F. (2026). Perceived quality of group activities fitness instructors: a comparative study across stakeholders, demographics, and club contexts. *Retos, 76*, 460-468.
https://doi.org/10.47197/retos.v76.117168

Abstract

Introduction: the perceived instructor quality (PIQ), in fitness group activities (GA), plays a vital role in users' satisfaction, motivation, and long-term engagement with fitness services.
Objective: this study aimed to assess and compare PIQ from both user and instructor perspectives, and to analyze how user perceptions vary according to gender, age group, and club attended.
Methodology: a total of 270 users and 32 instructors from five clubs of a fitness chain in Portugal completed an adapted version of the QIF-AG questionnaire, assessing 10 IQ items. Statistical analyses included independent samples t-tests and one-way ANOVA, with Tukey post hoc comparisons.
Results: results showed that users consistently rated instructors more positively than the instructors themselves, with statistically significant differences in their communication, punctuality, and knowledge. Female users reported higher IQ scores than males, particularly in relational dimension. While age-related differences were not statistically significant, younger users tended to give higher ratings across most items. Significant variations were also found among the clubs, some showing more favorable ratings, especially in technical-pedagogical dimension.
Discussion: findings underscored the multidimensional and context-sensitive nature of users PIQ. Fitness professionals and managers are encouraged to implement targeted quality improvement strategies, considering the user profiles and club characteristics.
Conclusions: this investigation highlights the value of integrating multiple perspectives and contextual variables when assessing service quality in fitness.

Keywords

Perceived instructor quality; fitness group activities; gender; group age; club attended.

Resumen

Introducción: la calidad percibida del instructor (CPI) en actividades grupales (AG) de fitness desempeña un papel clave en la satisfacción, la motivación y la adhesión a largo plazo de los usuarios.
Objetivo: evaluar y comparar la CPI desde las perspectivas de usuarios e instructores, y analizar cómo varía según el sexo, el grupo de edad y el club.
Metodología: participaron 270 usuarios y 32 instructores de cinco clubes de una cadena de fitness en Portugal, quienes completaron una versión adaptada del cuestionario QIF-AG, que evalúa 10 ítems de calidad del instructor. Se realizaron pruebas t para muestras independientes y ANOVA de un factor, con comparaciones post hoc de Tukey.
Resultados: los usuarios valoraron a los instructores de forma más positiva que la autoevaluación de los propios, con diferencias significativas en comunicación, puntualidad y conocimiento. Las mujeres obtuvieron puntuaciones de CPI más altas que los hombres, especialmente en la dimensión relacional. Aunque las diferencias por edad no fueron significativas, los usuarios más jóvenes tendieron a otorgar puntuaciones más elevadas. También se observaron variaciones entre clubes, sobre todo en la dimensión técnico-pedagógica.
Discusión: los resultados indicaron que la calidad percibida del instructor varía según la perspectiva de los participantes y el contexto del club, siendo especialmente relevantes los aspectos relacionales y técnico-pedagógicos.
Conclusiones: esta investigación ha mostrado la importancia de integrar múltiples perspectivas y variables contextuales en la evaluación de la calidad del servicio en fitness.

Palabras clave

Calidad percibida del instructor; actividades grupales de fitness; sexo; edad; club asistido.

Introduction

The fitness industry's growth (Franco et al., 2024; Newsome et al., 2024) has intensified the relevance of service quality as determinant of user satisfaction, loyalty, and organizational success (Braga-Pereira et al., 2025; Ferreira-Barbosa et al., 2022). Within fitness services, group activities (GA) are particularly popular (Palos et al., 2025; Pedragosa & Ferreira, 2025), and the role of GA instructors is essential for the users perceived instructor quality (PIQ) definition (Pedragosa & Ferreira, 2025). Essentially, instructors are the face and one of the most important variables of the service; they lead the class, engage with users, and directly shape the exercise experience (Pajaujiene, 2015; Silva et al., 2022). Recent research has consistently shown that instructor-related characteristics play a critical role in shaping users' PIQ; for instance, if GA instructors demonstrate competence, motivation, and effective communication skills, users are more likely to rate the class as high-quality and, consequently, the fitness center (Braga-Pereira et al., 2024; Glaveli et al., 2023; Xu et al., 2021).

Users PIQ, in GA, refers to the assessment of how well instructors deliver the class, both technically and interpersonally. It is a multidimensional construct, encompassing a range of behaviors and attributes. Campos et al. (2016a) developed *Qualidade do Instrutor de Fitness – Atividades de Grupo* (QIF-AG) scale [Quality of Fitness Instructor – Group Activities], to assess the PIQ in GA. This instrument identifies two PIQ dimensions: relational (RD); technical-pedagogical (TPD). RD includes interpersonal and affective characteristics (e.g., communication, good mood, sympathy, image, motivation) reflecting instructor's ability to connect with and inspire users. The TPD comprises skills related to the specificity of the area (e.g., instruction, planning, punctuality, knowledge, dedication). Together, the 10 skills/characteristics could capture, perceived by the users, different core facets of the GA instructor's performance.

It is crucial to recognize that PIQ may differ depending on observer's role (Campos et al., 2016b; Franco et al., 2013). Users form judgments based on experience and satisfaction during GA classes, whereas the instructors do their own performance self-assessment. Prior research suggests there can be meaningful gaps between the two perspectives. Franco et al. (2013) compared users' perceptions of GA instructor behavior with instructors self-perceptions, and found discrepancies across several behavioral items. The users and instructors do not always perceive it in the same way. A common pattern observed is that users often rate the instructors' performance more positively than instructors rate themselves.

Each user brings their unique profile, shaped by personal characteristics and differences (Gaspar, 2021). Individual characteristics can influence how they assess PIQ. Recent research highlights that perceived service quality in fitness settings can vary significantly with variables such as gender and/or age (Campos et al., 2021; Ortega Martínez et al., 2021). Tailoring communication style, class organization, and even technical guidance to different age groups and gender preferences, can enhance the PIQ. In fact, providing a more personalized approach in GA is likely to improve how quality is perceived by different groups of users (Campos et al., 2020; Glaveli et al., 2023). Such personalization strategy, adjusting the service to specific characteristics of distinct users' groups, have been recommended to boost PIQ, satisfaction, and loyalty, both in the service as a whole, and particularly in the GA (Braga-Pereira et al., 2024).

Beyond individual attributes, contextual factors, such as fitness club attended, can also impact users PIQ. Even within a fitness chain, each club may have a unique environment, users and management practices that could influence service execution (Amaral, 2025). The structural and organizational heterogeneity of fitness centers can directly affect the consistency and quality of the services provided, therefore shaping users' perceptions (Addolorato & Hormigón-Gimeno, 2025; Sadewa et al., 2025).

In summary, the literature suggests that users PIQ, in GA, is a multifaceted construction, influenced by instructors' behaviors, user characteristics, and contextual factors. Therefore, the present study aims to: (i) characterize and compare the PIQ in GA, from users and instructors' perspective; (ii) only considering the users perspective, characterize and compare the PIQ in GA, among gender, age group and club attended. By addressing both stakeholders' perspectives and these users' key variables, this study will contribute to a more nuanced understanding of users PIQ in fitness services, ultimately helping the professionals and managers develop targeted strategies for service personalization and improvement in both quality and users' satisfaction and loyalty (Eskiler & Safak, 2022; Xu et al., 2021).

Method

Participants

A non-probabilistic, convenience sampling strategy was employed, synchronized with the fitness chain management. All the GA instructors and users of the five clubs of a fitness chain in Portugal were invited. 270 users (U) and 32 instructors (I) participated, representing, respectively, a response rate of 62.79% and 94.12%. U are aged between 18 and 79 years old [Mean Standard Deviation (M±SD)=43.46±12.56], 224 (82.96%) are female (M±SD=42.33±12.68) and 46 (17.04%) male (M±SD=48.91±10.50). For I, ages ranges from 20 to 51 (M±SD=28.38±7.79), 13 (40.63%) female (M±SD=26.77±7.44) and 19 (59.37%) male (M±SD=29.47±8.04).

For U, the inclusion criteria were: 18 years old or over; having been enrolled in one of the five clubs for at least a month; and attending at least once a week. For I: at least 18 years old; working in one of the clubs; and having at least one month of professional experience at the club. All participants were informed about the subject of the study and filled out the informed consent form. The research was approved by Ethics Committee of the Polytechnic University of Coimbra (Portugal) (number D40/2024).

Instruments

For data collection, a reduced and adapted version of QIF-AG questionnaire, from Campos et al. (2016a), was used. The adaptation process included: (i) simplification of item wording for clearness; (ii) creation of two parallel versions [perception (Pv) for users, and self-perception (SPv) for instructors]. To assess users PIQ, 10 indicators were included (communication, good mood, sympathy, image, motivation: RD; instruction, planning, punctuality, knowledge, dedication: TPD), answered on a 7-point Likert scale from 1 (totally disagree) to 7 (totally agree). Higher values link to a better under assessment item. The content validity of this adapted version was reviewed by four experts in Sport Sciences, two in a first phase and two in the last, that considered it valid for the proposed objectives.

The reliability of the data was evaluated by its internal consistency, measured using Cronbach's alpha coefficient (Pallant, 2020). The scale achieved a α=.886 total coefficient and, for each dimension, α=.899 for RD and α=.873 for TPD. The alpha coefficients were classified as good (.8≤α<.9) (Pallant, 2020). The assessment of model fit quality was conducted through confirmatory factor analysis (CFA), based on the following fit indices: chi-squared ratio / degrees of freedom (χ^2/df), normal fit index (NFI), comparative fit index (CFI), Tucker-Lewis index (TLI), goodness-of-fit index (GFI), standardized root mean square residual (SRMR) (Hair et al., 2019; Marôco, 2021). The χ^2/df ratio, between 2 and 5 (3.513), indicates an acceptable fit; the values of NFI (.940), CFI (.948), TLI (.906), and GFI (.901), all between .90 and .95, indicate a good fit; additionally, SRMR value, below than .08 (.035), also indicate a good fit (Hair et al., 2019; Kline, 2023; Marôco, 2021).

Procedures

In a previous phase, collection process was articulated with the coordination of the fitness chain, which was in charge of framing the work and requesting the collaboration of all the instructors who provide service in the five clubs. GA users were recruited through in-class announcements made by instructors, complemented with follow-up invitations by the specific fitness club manager, via Salesforce's Customer Relationship Management (CRM) software.

Data collection was conducted online, through individualized survey links distributed by Salesforce's CRM. To maximize participation GA instructors encouraged users to contribute, emphasizing the study's aim of improving service quality. Data was collected from November 25 to December 20, 2024.

Statistical analysis

Descriptive statistics were used for characterization, specifically M±SD values. Independent t-test was used to compare perceived PIQ between stakeholders (instructors, users) and users' gender, after checking normality assumption (Field, 2028; Marôco, 2021). Normality assumption for each dependent variable was assessed using Kolmogorov–Smirnov test, for n≥30. In cases where the assumption of normality was not met, central limit theorem was applied (Field, 2018; Marôco, 2021). This theorem ensures that with sufficiently large samples (n≥30), the sampling distribution of means tends to be normal, allowing the use of the t-test even in the absence of data normality. Therefore, the normality assumption

was assumed to hold (Marôco, 2021; Pestana & Gageiro, 2014). When n<30, the assumption of normality for each dependent variable was assessed using Shapiro-Wilk test (Marôco, 2021; Pallant, 2020). In cases where normality was not verified, symmetrical analysis was performed using the following condition (Pestana & Gageiro, 2014; Tabachnick & Fidell, 2021):

$$\left|\frac{\text{skewness coefficient}}{\text{standard error of the skewness coefficient}}\right| \leq 1.96$$

One-way ANOVA was used to compare PIQ between users group age and club attended, after checking normality and homogeneity assumption (Pestana & Gageiro, 2014; Tabachnick & Fidell, 2021). The normality assumption is similarly assessed as for the independent t-test. To verify the assumption of homogeneity, the Levene statistical test was used (Field, 2018; Marôco, 2021). For multiple comparisons of means within groups, the Tukey HSD post hoc test was applied when the assumption of homogeneity was verified, while the Games-Howell test was used if the homogeneity assumption was not verified (Tabachnick & Fidell, 2021). For the data analysis, U were categorized into five age groups: <31, 31–40, 41–50, 51–60, and >60. The categories are based on previous studies in this area (Campos et al., 2016b), allowing their confrontation.

For independent t-test, the classification of the effect size (ES) is performed through reference values of Cohen's d (O'Donoghue, 2013): small (d<.20); moderate (.20≤d<.80); large (d≥.80). In turn, for one-way ANOVA, the ES classification was made according to the reference values of eta squared (η2) (Marôco, 2021): small (η2≤.05); moderate (.05< η2≤.25); high (.25< η2≤.50); very high (η2>.50). Apart from ES, the power (π) of corresponding test is also presented (Pallant, 2020). The statistical analysis was performed using IBM SPSS Statistics (version 28), for a significant level of 5% (p<.05).

Results

By the M±SD values is possible to verify that, in the 10 items, the PIQ in GA is higher in U, when compared to I self-perception (Table 1). Significant differences were found in three specific items [communication (t=2.700, p=.007, d=.505, moderate ES), punctuality (t=2.954, p=.006, d=.670, moderate ES), knowledge (t=2.305, p=.022, d=.431, moderate ES)], all presenting higher values for U. Cohen's d ES values reveal a moderate magnitude in the compared stakeholders PIQ.

Although other items did not present significant differences, U consistently rated I more positively across most indicators [good mood (t=1.857, p=.064, d=.347), planning (t=1.857, p=.064, d=.347), dedication (t=1.443, p=.150, d=.270)], with small-moderate ES. I only present higher results in three items [sympathy (t=-.138, p=.891, d=-.026), image (t=-.080, p=.936, d=-.015), motivation (t=.566, p=.572, d=.106)]. By dimension, there was a statistically significant difference in TPD (t=2.164, p=.031, d=.405, moderate ES), with U showing higher perceptions than I. In RD and total score (TS) there were no significant differences.

Table 1. Characterization and comparison of the users PIQ in GA, by stakeholder (users vs. instructors).

Item	Indicator	U	I	t	p	d
1.	Communication	6.62±1.01	6.13±.75	2.700	.007*	.505
4.	Good mood	6.51±1.12	6.13±1.01	1.857	.064	.347
5.	Sympathy	6.63±.92	6.67±.55	- .138	.891	- .026
6.	Image	6.74±.62	6.75±.57	- .080	.936	- .015
9.	Motivation	6.69±.88	6.59±.71	.566	.572	.106
	RD			1.322	.187	.247
2.	Instruction	6.46±1.18	6.28±.77	.846	.398	.158
3.	Planning	6.66±.94	6.34±.60	1.857	.064	.347
7.	Punctuality	6.85±.59	6.43±.76	2.954	.006*	.670
8.	Knowledge	6.72±.89	6.34±.79	2.305	.022*	.431
10.	Dedication	6.72±.85	6,50±.88	1.443	.150	.270
	TPD			2.164	.031*	.405
	TS			1.814	.071	.339

*significant for p < .05; U: Users (n=270), I: Instructors (n=32).

Considering the second objective (only considering the users' perspective, characterize and compare the PIQ in GA, among gender, age group and club attended), it is possible to observe the results in Table 2 (gender), 3 (group age) and 4 (club attended). Female users (S1) rated GA instructors more positively than male (S2) across all quality indicators under assessment (Table 2). Among these, significant differences were found in two items [good mood (t=2.035, p=.047, d=.437, moderate ES), punctuality (t=2.201, p=.032, d=.448, moderate ES)]. Like for stakeholder analysis, Cohen's d ES values confirm a moderate magnitude in PIQ differences, considering the users gender.

Table 2. Characterization and comparison of the users PIQ in GA, by users' gender.

Item	S1	S2	t	p	d
1.	6.67±.88	6.34±1.44	1.492	.142	.330
4.	6.59±.99	6.10±1.55	2.035	.047*	.437
5.	6.68±.87	6.36±1.10	1.841	.071	.347
6.	6.77±.57	6.56±.77	1.746	.086	.343
9.	6.73±.73	6.43±1.36	1.461	.150	.345
RD			1.901	.063	.423
2.	6.49±1.11	6.30±1.48	.997	.319	.161
3.	6.71±.78	6.39±1.46	1.451	.153	.347
7.	6.89±.54	6.63±.77	2.201	.032*	.448
8.	6.79±.69	6.36±1.46	1.922	.060	.486
10.	6.79±.66	6.43±1.40	1.673	.101	.424
TPD			1.637	.108	.408
TS			1.790	.079	.432

*significant for p < .05; S1: Female (n=224), S2: Male (n=47).

Analyzing users PIQ across different age groups, results show that overall scores remained consistently high among all age categories (Table 3). Users with 31-40 years (A2) generally assigned highest ratings, in nine of the 10 items. The five RD indicators showed a uniform pattern of positive assessments across all groups, with communication (A1: 6.78±.60; A2: 6.82±.52; A3: 6.65±1.07), image (A1: 6.80±.63; A2: 6.78±.58; A3: 6.80±.50) and motivation (A1: 6.70±.72; A2: 6.86±.40; A3: 6.65±.99) displaying the highest scores in younger users (A1–A3). However, differences were not significant. Similar trend was observed in TPD, where the higher scores were observed among the groups A2, A1 and A3, particularly in planning (A1: 6.80±.56; A2: 6.82±.52; A3: 6.62±1.15), punctuality (A1: 6.85±.56; A2: 6.94±.23; A3: 6.80±.77), and knowledge (A1: 6.85±.53; A2: 6.90±.46; A3: 6.65±1.14). Also, no statistically significant differences were found between groups.

Table 3. Characterization and comparison of the users PIQ in GA, by users group age.

Item	A1	A2	A3	A4	A5	F	p	η^2	π
1.	6.78±.60	6.82±.52	6.65±1.07	6.42±1.29	6.24±1.32	2.412	.059	.035	.689
4.	6.65±.91	6.74±.59	6.48±1.33	6.30±1.18	6.31±1.31	1.415	.229	.021	.438
5.	6.65±1.18	6.72±.57	6.59±.99	6.62±.85	6.58±.77	.166	.956	.002	.085
6.	6.80±.63	6.78±.58	6.80±.50	6.64±.69	6.55±.82	1.395	.236	.021	.432
9.	6.70±.72	6.86±.40	6.65±.99	6.62±1.00	6.55±1.05	.757	.554	.011	.242
RD						1.268	.283	.019	.355
2.	6.39±1.09	6.62±.87	6.55±1.16	6.26±1.46	6.41±1.32	.772	.554	.012	.247
3.	6.80±.56	6.82±.52	6.62±1.15	6.46±1.19	6.58±.77	1.306	.268	.019	.406
7.	6.85±.56	6.94±.23	6.80±.77	6.82±.59	6.86±.44	.445	.776	.007	.154
8.	6.85±.53	6.90±.46	6.65±1.14	6.56±1.05	6.55±.72	1.379	.242	.020	.427
10.	6.74±.71	6.88±.38	6.68±1.07	6.66±.98	6.68±.66	.541	.706	.008	.180
TPD						.904	.462	.013	.286
TS						1.090	.362	.016	.342

*significant for p < .05; A1: Less than 31 (n=54), A2: 31-40 (n=50), A3: 41-50 (n=87), A4: 51-60 (n=50), A5: More than 60 (n=29).

When comparing the users PIQ according to the club attended, results show that users from C3 and C5 generally attributed the highest scores across most indicators, while C2 consistently presented lower ratings (Table 4). For instance, users from C3 assigned high ratings to planning (6.80±0.42), knowledge (6.81±0.56), and motivation (6.85±0.41), while C5 stood out particularly in dedication (6.84±0.46). In contrast, participants from C2 gave the lowest ratings to several indicators, such as communication (6.23±1.57), motivation (6.34±1.52), instruction (6.04±1.69), planning (6.23±1.59), knowledge

(6.34±1.59), and dedication (6.36±1.52). Significant differences were observed in six indicators: communication (F=2.639, p=.034, η^2=.038, small ES), with users from C5 (6.78±.74) and C3 (6.75±.62) presenting higher PIQ than C2 (6.23±1.57); motivation (F=2.819, p=.026, η^2=.041, small ES), with significant differences between C3 (6.85±.41) and C2 (6.34±1.52); instruction (F=3.081, p=.017, η^2=.044, small ES), with significant differences between C3 (6.66±.62) and C2 (6.04±1.69); planning, (F=3.195, p=.014, η^2=.046, small ES), also with differences between C3 (6.80±.22) and C2 (6.23±1.59); knowledge (F=2.707, p=.031, η^2=.039, small ES), with significant differences between C3 (6.81±.56) and C2 (6.34±1.59); dedication (F=2.804, p=.026, η^2=.026, small ES), C5 (6.84±.46) and C3 (6.80±.60) presenting higher IQ than C2 (6.36±1.52). There were also significant differences in TPD (F=2.907, p=.022, η^2=.042, small ES), from C3 and C2 users. The small ES values don't allow confirmation that there are substantial differences in PIQ, considering the users' club attended. No differences were found in the RD (F=1.585, p=.179, η^2=.023) or in TS (F=2.263, p=.063, η^2=.033) although the trend suggests that C3 and C5 users assume the PIQ more positively.

Table 4. Characterization and comparison of the users PIQ in GA, by users club attended.

Item	C1	C2	C3	C4	C5	F	p	η^2	π
1.	6.66±.73	6.23±1.57[a,c]	6.75±.62[a]	6.54±1.17	6.78±0.74[c]	2.639	.034*	.038	.754
4.	6.45±1.09	6.29±1.57	6.67±.67	6.52±1.10	6.44±1.27	.977	.420	.015	.308
5.	6.60±1.19	6.59±.97	6.72±.75	6.72±.63	6.44±1.14	.922	.452	.014	.291
6.	6.69±.63	6.59±.77	6.81±.51	6.76±.61	6.75±.62	1.023	.396	.015	.322
9.	6.60±.70	6.34±1.52[a]	6.85±.41[a]	6.76±.81	6.69±.75	2.819	.026*	.041	.775
RD						1.585	.179	.023	.486
2.	6.63±.74	6.04±1.69[a]	6.66±.62[a]	6.23±1.60	6.61±.97	3.081	.017*	.044	.806
3.	6.75±.50	6.23±1.59[a]	6.80±.42[a]	6.72±.89	6.67±.94	3.195	.014*	.046	.822
7.	6.84±.44	6.76±.66	6.88±.44	6.96±.19	6.75±.96	1.126	.345	.017	.353
8.	6.78±.48	6.34±1.59[a]	6.81±.56[a]	6.80±.60	6.78±.82	2.707	.031*	.039	.764
10.	6.81±.46	6.36±1.52[a,b]	6.80±.60[a]	6.84±.46[b]	6.76±.75	2.804	.026*	.041	.773
TPD		[a]	[a]			2.907	.022*	.042	.790
TS						2.263	.063	.033	.657

*significant for p < .05; C1: Club 1 (n=33), C2: Club 2 (n=47), C3: Club 3 (n=87), C4: Club 4 (n=51), C5: Club 5 (n=52); multiple comparison using post-hoc test: a) C2 vs. C3, b) C2 vs. C4, c) C2 vs. C5.

Discussion

Findings of this research reinforce the relevance of GA PIQ, in fitness centers, highlighting the impact of stakeholders' role, as well as sociodemographic and contextual variables. Consistent with previous research, users tended to rate instructor performance more positively than instructors rated themselves (Franco et al., 2013; Campos et al., 2016b), with significant differences found in specific items such as communication, punctuality, and knowledge. This pattern supports the concept that users could assess quality based on their own lived experience, often valuing affective and relational aspects (Campos et al., 2020; Gaspar, 2021), whereas instructors may adopt a more self-critical stance, emphasizing some technical and pedagogical standards (Campos et al., 2016b; Franco et al., 2013). Significant differences observed in TPD strengthens this, since the dimension includes more objectively assessable aspects of professional practice. This asymmetric perception likely arises from role-dependent evaluative criteria: users' privilege experiential and relational cues during the service encounter; GA instructors benchmark themselves against technical standards and pedagogical routines. The fact that differences concentrate in TPD suggests a calibration gap on "how" classes are performed (e.g., punctuality) rather than "who" the instructor is, pointing to the value of structured feedback and peer observation focused on technical-pedagogical execution (Braga-Pereira et al., 2024; Glaveli et al., 2023).

Regarding gender variable, female users consistently perceived higher PIQ, particularly in the items of good mood and punctuality. While the results may partly reflect the sample gender distribution (mostly female), it also aligns with previous evidence, suggesting that women may be more sensitive to interpersonal dynamics and relational quality in service contexts (Campos et al., 2021; Ortega-Martínez et al., 2021). Female users may value emotional engagement and communication to a greater extent, thus reporting more favorable perceptions of GA instructor behaviors. Two non-mutually exclusive mechanisms may explain this pattern: (i) the women may assign greater weight to relational quality (e.g., empathy) when forming overall judgments (Ortega-Martínez et al., 2021; Campos et al., 2021); (ii) exposure differences [women attend GA formats where choreographic guidance and continuous instructor

communication is central (Pajaujiene, 2015)], can amplify the salience of this cue. Practically, emphasizing relational these micro-skills (e.g., clarifying cues, positive affect, timely feedback) could yield disproportionate benefits for female users.

Although no significant differences were found between age groups, younger users generally reported higher scores across quality indicators under analysis. This trend, although not conclusive, suggests that age may influence expectations and interpretations of the users PIQ, potentially linked to differences in communication style, exercise intensity, or instructor approach (Gaspar, 2021; Silva et al., 2022). These findings are in line with recommendations for personalization strategies, tailored to user characteristics (Braga-Pereira et al., 2024; Glaveli et al., 2023). Two factors can help interpret these age trends: (i) a ceiling effect (high means with limited variance reduce detectable differences); (ii) cohort-linked expectations (younger users may value immediacy, digital integration, and energetic pacing, which align with higher ratings for instruction and planning) (Gaspar, 2021; Silva et al., 2022).

The differences observed between clubs further highlight the role of contextual factors in shaping PIQ. Users from C3 and C5 rated PIQ with higher values across several indicators, while C2 users consistently reported the lowest scores. These differences confirm that even within the same fitness chain, local club dynamics, such as environment, management and organizational culture, may influence service execution and clients' experience (Addolorato & Hormigón-Gimeno, 2025; Sadewa et al., 2025). This reinforces the importance of context-sensitive quality management strategies, capable of adapting to each location's characteristics. The pattern is consistent with contextual drivers – service climate, scheduling discipline, staff stability, workload, and facility ambience – that shape the same instructor's performance envelope (Addolorato & Hormigón-Gimeno, 2025; Sadewa et al., 2025). Management (e.g., timetable reliability, class density, briefing routine) are therefore plausible targets to reduce the between-club dispersion in PIQ.

The result of this research supports the conceptualization of the GA users PIQ as a multidimensional construct, influenced by individual, interpersonal and contextual factors, as emphasized in the literature (Braga-Pereira et al., 2024; Glaveli et al., 2023). For effective quality management in the fitness services, professionals and managers must understand these specific nuances and design targeted interventions that reflect user profiles and local contexts. This may be the key that helps fitness centers to differentiate the services and retain the users, thus being able to succeed in an increasingly competitive market, an idea corroborated by Amaral (2025) and Jaques (2025).

Conclusions

Results are consistent with the literature, showing that users tend to rate PIQ higher than instructors themselves, especially in technical-pedagogical dimensions. Gender-based differences suggest that female U may perceive instructors more positively, particularly in relational aspects, while age-related trends, although not statistically significant, indicated that younger U might attribute higher ratings across most PIQ indicators. Furthermore, club-based comparison suggested that the context in which the service is delivered may influence how users understand the PIQ, although these contextual differences should be interpreted with caution.

These findings have practical implications. For GA instructors, the results emphasize the importance of instructors improving their performance considering users perception, while promoting self-awareness and reflective practices. For instructors and fitness club managers, understanding the possible variations in PIQ across demographic groups and club environments may support the development of targeted strategies to enhance user satisfaction, engagement, and loyalty, key factors in a highly competitive market.

Limitations and recommendations

Despite its contributions, this study has some limitations. First, the sample of instructors was relatively small and unbalanced by gender, which restricts generalizability. Second, the study relied on self-reported measures, which may be influenced by social desirability or individual bias. Third, although the

study explored contextual club-level differences, it did not account for deeper organizational variables such as management style, instructor workload, or offered GA. Future research should consider expanding the instructors' sample, integrating observational and qualitative methods to enrich the understanding of users PIQ. It may also be relevant to explore longitudinal PIQ effects on users' satisfaction or loyalty, or to test intervention strategies aimed at improving PIQ in specific users' segments or fitness centers contexts.

Financing

FCT – Fundação para a Ciência e a Tecnologia – Foundation for Science and Technology, I.P. (Portugal), within the scope of SPRINT – Sport Physical Activity and Health Research & Innovation Center [UID/6185/2023].

References

Addolorato, S., & Hormigón-Gimeno, N. (2025). An analysis of active consumer behavior in a fitness boutique: Satisfaction levels, perceived quality dimensions, and dropout reasons. *Retos*, *62*, 362–373. https://doi.org/10.47197/retos.v62.109312

Amaral, A. (2025). A fotografia menos bonita dos clubes de fitness e saúde em Portugal. In Portugal Activo (Ed.), *Sete contributos para partilhar interpares* (pp. 11–24). Portugal Activo.

Braga-Pereira, R., Furtado, G., Campos, F., Sampaio, A., & Teques, P. (2024). Impact of fitness coach behavior on exercise motivation, commitment, and enjoyment: A longitudinal study. *PLOS One, 19*(12), e0310931. https://doi.org/10.1371/journal.pone.0310931

Braga-Pereira, R., Furtado, G., Campos, F., Sampaio, A., & Teques, P. (2025). Enhancing exercisers' experience: A comprehensive review. *International Journal of Sport and Exercise Psychology*, 1–22. https://doi.org/10.1080/1612197X.2024.2437480

Campos, F., Simões, V., & Franco, S. (2016a). A qualidade em atividades de grupo de fitness: Construção e validação do questionário "Qualidade do Instrutor de Fitness-Atividades de Grupo" (QIF-AG). *Psicologia*, *30*(1), 37–48. https://doi.org/10.17575/rpsicol.v30i1.1069

Campos, F., Simões, V., & Franco, S. (2016b). Characterization and comparison of the quality indicators of group exercise fitness instructor, considering the intervenient, gender and age. *International Journal of Sports Exercise Training Science*, *2*(2), 50–59. https://doi.org/10.18826/ijsets.63014

Campos, F., Simões, V., & Franco, S. (2020). A qualidade do instrutor em atividades de grupo de fitness. In S. Franco & V. Simões (Eds.), *Pedagogia do fitness. Contributos para a intervenção dos profissionais* (pp. 7–29). Omniserviços.

Campos, F., Martins, F., Gomes, R., Zylberberg, T., Mendes, S., & Damásio, A. (2021). Importância atribuída no fitness segundo o rendimento líquido mensal dos praticantes. *Retos*, *40*, 336–343. https://doi.org/10.47197/retos.v1i40.81957

Eskiler, E., & Safak, F. (2022). Effect of Customer Experience Quality on Loyalty in Fitness Services. Physical Culture and Sport. Studies and Research, 94, 21-34. https://doi.org/10.2478/pcssr-2022-0003

Ferreira-Barbosa, H., Barbosa, J., Sabino, B., & Loureiro, V. (2022). Determinants of service quality influencing customer satisfaction in fitness centers: A systematic review. *European Journal of Human Movement*, *49*, 29–45. https://doi.org/10.21134/eurjhm.2022.49.3

Field, A. (2018). *Discovering statistics using IBM SPSS statistics*. SAGE.

Franco, S., Simões, V., Castañer, M., Rodrigues, J., & Anguera, M. (2013). La conducta de los instructores de fitness: Triangulación entre percepción de los practicantes, auto-percepción de instructores y conducta observada. *Revista de Psicología del Deporte*, *22*(2), 321–329.

Franco, S., Simões, V., Santos-Rocha, R., Vieira, I., Ramalho, F., & Ramos, L. (2024). Tendências do fitness em Portugal para 2024. *Retos*, *57*, 88–100. https://doi.org/10.47197/retos.v57.105198

Gaspar, C. (2021). Metodologia das atividades de grupo. In D. Teixeira, P. Marques, & C. Gaspar (Eds.), *Manual do técnico de exercício físico* (pp. 253–292). MANZ.

Glaveli, N., Papadimitriou, D., Karagiorgos, T., & Alexandris, K. (2023). Exploring the role of fitness instructors' interaction quality skills in building customer trust. *European Sport Management Quarterly*, *23*(3), 767–788. https://doi.org/10.1080/16184742.2021.1928256

Hair, J., Black, W., Babin, B., & Anderson, R. (2019). *Multivariate data analysis*. Cengage Learning.

Jaques, B. (2025). Marketing estratégico. In Portugal Activo (Ed.), *Sete contributos para partilhar* (pp. 27–42). Portugal Activo.

Kline, R. (2023). *Principles and practice of structural equation modeling*. Guilford Publications.

Marôco, J. (2021). *Análise estatística com o SPSS statistics*. Report Number.

Newsome, A., Batrakoulis, A., Camhi, S., McAvoy, C., Sansone, J., & Reed, R. (2024). 2025 ACSM worldwide fitness trends: Future directions of the health and fitness industry. *ACSM's Health & Fitness Journal*, *28*(6), 11–25. https://doi.org/10.1249/FIT.0000000000001017

O'Donoghue, P. (2013). *Statistics for sport and exercise studies: An introduction*. Routledge.

Ortega-Martínez, J., Martins, F., González-Víllora, S., & Campos, F. (2021). Estudio sobre la calidad percibida de los usuarios de fitness: A partir de sus características y preferencias según sexo y edad. *Retos, 39*, 477–482. https://doi.org/10.47197/retos.v0i39.80061

Pajaujiene, S. (2015). Delivering a group fitness class. In R. Santos-Rocha, T. Rieger, & A. Jiménez (Eds.), *Europeactive's essentials for fitness instructors* (pp. 85–104). Human Kinetics.

Pallant, J. (2020). *SPSS survival manual: A step-by-step guide to data analysis using IBM SPSS*. Routledge.

Palos, J., Valcarce-Torrente, M., Amaral, P., Franco, S., Santos-Rocha, R., Ramalho, F., Simões, V., Vieira, I., Ramos, L., Chávez, L., Almanzar, P., López-Haro, J., Zavalza, A., Rodrígez, L., Giakoni-Ramírez, F., Muñoz-Strale, C., Souza-Lima, J., Pinochet-Peña, F., Yáñez-Sepúlveda, R., …Veiga, O. (2025). Iberoamerican fitness trends for 2025: A regional comparative study. *Retos, 71*, 1095–1114. https://doi.org/10.47197/retos.v71.116451

Pedragosa, V., & Ferreira, A. (2025). *Barómetro do fitness Portugal Activo 2024*. Portugal Activo.

Pestana, M., & Gageiro, J. (2014). *Análise de dados para ciências sociais: A complementaridade do SPSS*. Sílabo.

Sadewa, Y., Sumaryanto, S., & Sumarjo, S. (2025). Marketing strategies to improve service quality at sports and fitness venues. *Retos*, *62*, 797–801. https://doi.org/10.47197/retos.v62.109886

Silva, A., Brás, D., Almeida, H., & Máximo, O. (2022). Metodologia das atividades de grupo. In C. Olim, H. Zebre, & R. Gonçalves (Eds.), *Manual do técnico especialista em exercício físico* (pp. 351–367). Omniserviços.

Tabachnick, B., & Fidell, L. (2021). *Using multivariate statistics*. Pearson.

Xu, K., Chen, K., Kim, E., García-Fernández, J., Nauright, J., & Zhang, J. (2021). Dimensions of Service Quality in Health-Fitness Clubs in China. *International Journal of Environmental Research and Public Health*, *18*(20), 10567. https://doi.org/10.3390/ijerph182010567

Authors and translators' details:

Daniela Almeida	danssousa99@hotmail.com	Author
Vera Simões	verasimoes@esdrm.ipsantarem.pt	Author
Susana Franco	sfranco@esdrm.ipsantarem.pt	Author
José Aleixo	aleixo.uc@gmail.com	Author
Fernando Martins	fmlmartins@esec.pt	Author
Francisco Campos	francicampos@esec.pt	Author