Effect of gamification-based programs in the school physical education setting: a systematic review protocol

Efecto de los programas basados en gamificación en el contexto de la Educación Física escolar: Un protocolo de revisión sistemática

Carlos Alberto Becerra-Fernández[1]

[1]Departamento de Didáctica de las Lenguas, las Artes y el Deporte, Facultad de Ciencias de la Educación, Universidad de Málaga, Málaga, España

Abstract

Introduction: Gamification is an innovative methodology that uses game elements, such as, for instance, narrative, rewards, or avatars, to influence student´s motivation with the aim of improving the teaching and learning process. The main purpose of the systematic review presented in this protocol will be to examine the scientific literature on the effects of gamified didactic proposals in the context of Physical Education on learning and its mediators in schoolchildren. **Method:** The databases search following the search strategies and the download and collection in the reference manager will be completed from its insertion to December 2022. This search will include the following electronic bibliographic databases: Web of Science[TM] (all databases), Scopus, SPORTDiscus with Full Text, ERIC, and ProQuest Dissertations & Theses Global[TM]. **Results:** This systematic review will be a useful resource for teachers and researchers belonging to the area from Physical Education, since they will be able to easily know the current evidence about the topic, to know how this methodology is designed and its effectiveness compared to traditional methodologies. **Discussion and conclusions:** The systematic review to be carried out in the near future can provide information on the recommended length of gamification programs in Physical Education, its unique elements for improve the teaching and learning process, as well as describe the impact of intrinsic motivation on academic performance, improve the participation and increase of physical activity levels.

Key words: Gamified, motivation, pedagogical model, active methodology, schoolchildren.

Resumen

Introducción: La gamificación es una metodología innovadora que utiliza elementos del juego, como por ejemplo, la narrativa, recompensas o avatares, para influir en la motivación de los estudiantes con el objetivo de mejorar el proceso de enseñanza y aprendizaje. El objetivo principal de la revisión sistemática presentada en este protocolo será el de examinar la literatura científica sobre los efectos de las propuestas didácticas gamificadas en el contexto de la Educación Física sobre el aprendizaje y sus mediadores en los escolares. **Método:** Se completará la búsqueda en bases de datos siguiendo las estrategias de búsqueda, descarga y recopilación en el gestor de referencias, desde su inserción hasta diciembre de 2022. Esta búsqueda incluirá las siguientes bases de datos bibliográficas electrónicas: Web of ScienceTM (todas las bases de datos), Scopus, SPORTDiscus with Full Text, ERIC y ProQuest Dissertations & Theses GlobalTM. **Resultados:** Esta revisión sistemática será un recurso útil para docentes e investigadores pertenecientes al área de Educación Física, ya que podrán conocer con facilidad la evidencia actual sobre el tema, para saber cómo está diseñada esta metodología y su efectividad frente a las metodologías tradicionales. **Discusión y conclusiones:** La revisión sistemática llevada a cabo en un futuro próximo puede aportar información sobre la duración recomendada de los programas de gamificación en la Educación Física, sus elementos fundamentales para mejorar el proceso de enseñanza y aprendizaje, así como describir el impacto que tiene la motivación intrínseca sobre el rendimiento académico, la participación y el incremento de los niveles de actividad física.

Palabras claves: Gamificado, motivación, modelo pedagógico, metodología activa, escolares.

INTRODUCCIÓN

According to Fernández-Río and Flores Aguilar (2019), the Gamification can be understood as a pedagogical model that uses the elements of the game (narrative, aesthetics, prizes) to develop concrete curricular content within a context which includes tasks and activities adapted to the dynamics of the game, to achieve the educational objectives set and not the simple deviation. Werbach and Hunter (2012) affirm that it should take into consideration three large categories in all gamified context: Dynamics, mechanics and components. Within each are numerous key elements such as narrative, challenges, learning cooperative, avatars, badges, points, levels and rankings. These facilitate the teaching-learning process and the connections between students and with the teacher (Hanus& Fox, 2015).

The application of this type of active methodology is essential in teaching, because demotivation and lack of interest in learning are common problems in the educational field. Therefore, gamification is considered an educational important model which offers positive experiences and favors a more significant learning among the students, meaning that an improvement in their basic psychological needs and level of motivation.

For example, the inclusion criteria of other studies only select them from 2015 (Piñeiro Aboy, 2022) and others from 2010 onwards (Hernández et al., 2021). In other cases, they limit the language to Spanish and English (Hernández et al., 2021; Piñeiro Aboy, 2022). They even carry out a very general search that includes any aspect of gamification in Physical Education (Escaravajal & Martín-Acosta, 2019). Specifically, this last author does not specify how the research design should be, whether there should be a control group or not and if the

studies that also use exergames would enter.

One of the limitations of these reviews is that they do not establish in their selection criteria if the works are exclusively experimental or quasi-experimental designs, that is, they have at least a control group and/or baseline measurement in order to compare the dependent variables before and after the intervention. A systematic review that provides current and robust evidence about the effects of Gamification in the Physical Education setting is needed. Consequently, the main purpose of the systematic review presented in this protocol will be to examine the scientific literature on the effects of gamified didactic proposals in the context of Physical Education on learning and its mediators in schoolchildren.

METHODS

The present systematic review will be based on the methodology described in previous reference literature such as the PRISMA guidelines (Page et al., 2020) and Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2021), among other important references (Borenstein et al., 2009; Cooper et al., 2019; Casado-Robles et al., 2022).Firstly, a reproducible, transparent, and comprehensive systematic review will be performed to identify, select, and synthesize all the relevant studies. Then, in case that there will be enough number of similar studies with the same dependent variable/s, a meta-analysis will be performed to provide more precise estimates of the effects than those derived from the primary studies.

Eligibility criteria

The eligibility criteria for including the retrieved studies in the systematic review will be the following: (1)Participants: apparently healthy schoolchildren (5-17 years old); (2) Intervention: studies that examine the effect of gamified programs (not exclusively exergames) on schoolchildren´s learning and its mediators (both sessions and teaching units are included); (3) Comparator: They must include at least one measure to compare intervention effects(i.e., baseline measurements in single-group designs and/or a control group with no intervention or with traditional intervention). Thus, interventions with only one experimental group and only post-intervention measures were not included; (4)Outcomes: studies that assess the effect of the gamified programs on the schoolchildren´s learning and/or its mediators; (5) Study design: Any kind of experimental designs including, but not limited to, pre-experimental trials (e.g., non-controlled trials with one-group pre–post-intervention design); quasi-experimental trials (e.g., cluster-randomized controlled trials or non-randomized controlled trials), and true-experimental trials (e.g., randomized controlled trials) will be included for examining the effect of a gamification-based program such as a teaching unit or similar. Moreover, it will be selected acute effect studies where a gamification-based session or similar is compared with another type of session (e.g., cross-over or cluster-cross over); (6) Setting: The gamified program is carried out exclusively, or at least in part, in the context of school Physical Education.

Data sources and search strategy

The databases search following the search strategies and the download and collection in the reference manager will be completed from its insertion to December 2022. This search will include the following electronic bibliographic databases: Web of Science™ (all

databases), Scopus, SPORTDiscus with Full Text, ERIC, and ProQuest Dissertations & Theses Global^TM. The searches will carry out in the search field type "title, abstract, and keywords" or equivalent. The search terms used will be based on two concepts: (1) Gamification, and (2) Physical Education setting. The terms of the same concept will be combined with the Boolean operator "OR" and then the two concepts will be combined using the Boolean operator "AND." The keywords with more than one word will be enclosed in quotes. No publication status, language, or date restrictions will be imposed (Cooper et al., 2019).Specifically, the syntax used in all databases will be the next one: (gamification OR ludification OR gamified OR exergame OR exergames OR "active methodology" OR "active methodologies") AND ("physical education*" OR "PE" OR "P.E." OR school*).

Then, additional studies will be identified as follows: (1) searching the reference lists of original studies, as well as some related study reviews and study protocols; (2) examining the reference citations and the researchers' publications (first authors) in the Web of Science^TM and Scopus databases; (3) contacting with the corresponding authors by email, and (4) screening there searchers' personal lists in ResearchGate and GoogleScholar (first authors). Any time a new study is found, all of these modes of searching will be repeated until nonew study appears.

Study selection

After eliminating duplicates, the first author will undertake the potentially eligible records selection based on the screening of titles and abstracts to identify relevant studies. After that, two independent reviewers will assess the full texts for inclusion following the above-mentioned eligibility criteria. Any disagreements regarding the inclusion of studies will be resolved by consensus with a third reviewer.

Data extraction

From each selected study, data will be coded using an *ad hoc* coding form developed by the research group and previously tested with a pilot sample of studies at the beginning of the review. This form will be included data about:(1) study characteristics (i.e., reference, publication date, date of the data collection, study design, suspicion of selective outcomes, and initial and final sample size); (2) participant characteristics (i.e., sex and age); (3) outcome measures pertaining to dependent variables (i.e., measurement moment, measurement time, kind of measurement instrument, and measurement score and units); (4) intervention characteristics(intervention characteristics and length); and(5) results of the intervention for each group (i.e., initial and final group size, pre- and post-intervention standard deviation, and pre- and post-intervention means score or pre–post-intervention mean difference score).

If a study consists of two or more study arms of which one of the intervention arms do not meet the inclusion criteria, data will be only extracted from the study arms that met the inclusion criteria. In the event that the studies do not report some study feature, corresponding authors will be contacted to retrieve it. If means and standard deviation will not be retrieved, the scores will be estimated and converted by the standard error, confidence intervals, F, t or p values (Higgins et al., 2021). Since median and inter quartile range are often used when the data are asymmetrical, these values will not be converted (Higgins et al., 2021). If any other study feature is not retrieved, the information will be omitted. The sample size of each group, the mean scores of the pre- and/or post-intervention or mean difference scores of each group,

and the measurement score of the dependent variable/s will be considered to be critical for including the selected studies from the systematic review in the possible meta-analysis. In order to avoid removing studies from the meta-analysis, numerical data will be extracted from their figures using the WebPlotDigitizer software (Higgins et al., 2021). Coding studies will be carried out independently by two researchers. When doubt or disagreement occurs, a consensus will be always achieved through discussion.

Risk of bias

Based on the Cochrane risk-of-bias tool version 2 (Higgins et al., 2021),the following methodological domains will be assessed:(1) risk of bias arising from the randomization process; (2) risk of bias due to missing outcome data; (3) risk of bias in measurement of the outcomes, and (4) risk of bias in the selection of the reported results. Due to the nature of the selected studies, some risk of bias criteria will be not considered. Firstly, blinding of participants and personnel will be not included due to participants being always aware of them. Therefore, the assessment of allocation concealment will be also meaningless in this type of study since both participants and researchers would know which group they were assigned to during the intervention phase. Finally, in case of objective measurements, blinding the outcomes assessment will be not applicable. Domains will be judged and classified as "low risk,""some concerns" or "high risk" of bias. Finally, over-all risk of bias judgment will be obtained as follows: "low risk" if low risk of bias is obtained for all domains; "some concerns" if at least one domain is judged as having some concerns, but not to be at high risk of bias for any domain; and "high risk" if at least one domain is judged as high risk or if two or more domains are judged as having some concerns (Higgins et al., 2021).

Certainty of the evidence

In case that there will be enough number of similar studies with the same dependent variable/s, the overall certainty of the evidence will be rated as "high," "moderate," "low," or "very low" using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach (Guyatt et al., 2011). This assessment will be based on the following five domains: risk of bias, inconsistency, imprecision, indirectness, and publication bias. A domain will be classified as "no limitation" if no reason for downgrading the evidence is found, but the domain will be classified as "serious" if a reason is found for downgrading the evidence(thus, downgrading the certainty rating by one level).The reasons for considering the domains as "serious" are: (1) Risk of bias: Most of the studies (i.e., $> 50\%$)are classified as "high" risk of bias; (2) Inconsistency: high level of heterogeneity (i.e., $I^2 > 75\%$) is found; (3)Imprecision: The confidence interval will wide including both the possibility of trivial effect (i.e., $d < 0.20$)and large effect (i.e., $d \geq 0.80$); (4) Indirectness: Most studies (i.e., $> 50\%$) address a restricted version of the main review question in terms of population, intervention, comparator or outcomes; and (5) Publication bias: Egger's test is statistically significant and the impact of publication bias is large (i.e., the number of additional studies with no effect that would be needed to increase the p value over a statistically insignificant effect (i.e., $p > 0.05$) is low following the fail-safe N analyses; or Trim and Fill method trimmed many studies with an adjusted value for effect size different to the observed values).

Data analyses

In case that there will be enough number of similar studies with the same dependent variable/s, meta-analysis will be performed in order to provide more precise estimates of the effects than those derived from the primary studies. Particularly, meta-analysis will include the following issues: (1) Effects sizes computation: Meta-analyses of the Cohen's d standardized mean difference and 95% confidence interval with a random-effects model will be conducted to obtain the intervention program effects of dependent variables; (2) Publication bias: Firstly, an exhaustive systematic review will carry out to avoid availability bias. Afterward, a deep examination of the selected studies will carry out to avoid any potential duplication of the information retrieved. Then, to identify the impact of any potential publication bias, the funnel plots and the Egger's test (Egger et al., 1997) will carry out. Moreover, for assessing the impact of any potential publication bias, the Orwin's fail-safe N analyses and the Duval and Tweedie's Trim and Fill method (Duval & Tweedie, 2000) will be computed; (3)Heterogeneity: The presence of statistical heterogeneity in the estimation of the effect sizes of the program will examine with the I^2statistic; (4)Subgroups analyses: Based on a priori hypothesized moderators, subgroups analyses will also carry out to test the effect of the gamified programs regarding educational stage (i.e., Primary and Secondary Education), and (5) Sensitivity analyses: Cohen's *d* with a fixed-effect model, Hedges' *g* with a random-effects model, and Cohen's *d* with a random-effects model separately for (cluster-) randomized controlled trial design or not. A detailed description of how the data analyses will be carried out in the present meta-analysis can be found elsewhere (Casado-Robles et al., 2022).The meta-analyses will be performed using the software current version of the Comprehensive Meta-Analysis for Windows (Biostat, Englewood, USA).

DISCUSSION AND CONCLUSIONS

Previous systematic reviews about the effects of Gamification in the Physical Education setting have shown some limits regards to design of studies, some of them, without control group or with an only one post-intervention measure. Even there are reviews that include gamified programs in different contexts and select interventions based on exergames which could create confuse.

One of the aims of the present systematic review protocol is to explain these limitations that have not been very clarify in previous studies. Specifically, the systematic review will investigate the association between different dependent variables (for example, intrinsic motivation, basic psychological needs, academic performance, or level of physical activity) belonging to the same context, in this case, Physical Education lessons. Not only will be clarify the dependent variables most observed and measured in this context, but also the assessment instruments most used. Clarifying these elements will make this review have a good scientific consistency, since it will provide a detailed guide for Physical Education teaching professionals and academics to promote and prescribe gamified programs adequately, knowing their elements, measurement instruments and dependent variables in depth.

Strengths of the systematic review that it is described in the present protocol are that it will guide the authors to conduct a valid and reliable systematic review considering all factors related to the risk of bias, study eligibility, comprehensive search strategy, and the inclusion

of psychological outcomes such as basic psychological needs and motivation in learning. Performing an inclusive search in major databases as those reported previously also will also some major strengths to the systematic review to be conducted. Results of the present study are expected to provide a deeper understanding of the difference between gamified methodology and exergames in education and how a gamification program can affect on motivation and physical activity levels.

This protocol provides the justification and planning methods for a systematic review, aiming to analyze the impact of Physical Education interventions based on gamified sessions and teaching units in schoolchildren. We will consider the strengths and limitations of the studies identified in the search procedure, as well as those of our proposed systematic review, and discuss the results in the context under analysis.

The systematic review to be carried out in the near future can provide information on the recommended length of gamification programs in Physical Education, its unique elements for improve the teaching and learning process, as well as describe the impact of intrinsic motivation on academic performance, improve the participation and increase of physical activity levels. What it means, motivation is essential to change inappropriate behaviors for others which be much healthier with the objective of improving the quality of life of students.

The results of the systematic review would help professionals and scholars involved in the Physical Education-based programs to develop and implement gamified interventions that effectively promote a healthy behavior related to physical activity and care of emotions in schoolchildren. As a consequence, it is theoretically expected that students with demotivation, bad academic results, anxiety, sedentary habits and low participation levels, could foster positive outcomes during and after a gamified intervention increasing physical activity, participation, autonomy, motivation, quality of life and subjective well-being. In order to disseminate the proposed systematic review, the results will be published across academic publications, conferences, peer-reviewed presentations, and formal meetings.

REFERENCES

Borenstein, M. M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introductionto meta-analysis*. 2nd ed. Wiley.

Casado-Robles, C., Viciana, J., Guijarro-Romero, S., & Mayorga-Vega, D. (2022). Effects of consumer-wearable activity tracker-based programs on objectively measured daily physical activity and sedentary behavior among school-aged children: A systematic review and meta-analysis. *Sports Medicine - Open*, *8*(1), 18. https://doi.org/10.1186/s40798-021-00407-6

Cooper, H., Hedges, L. V., & Valentine, J. C. (2019). *The handbook of research synthesis and meta-analysis.* 3rd ed. Russell Sage Foundation.

Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455-63. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629-634. https://**10.1136/bmj.315.7109.629**

Escaravajal, J. C., & Martín, F. (2019). Análisis bibliográfico de la gamificación en Educación Física. *Revista Iberoamericana de Ciencias de la Actividad Física y del Deporte, 8*(1), 97-109. https://doi.org/10.24310/riccafd.2019.v8i1.5770

Fernández-Río, J., & Flores-Aguilar, G. (2019). Fundamentación teórica de la Gamificación. En J. Fernández-Río (Coord.), *Gamificando la Educación Física. De la teoría a la práctica en Educación Primaria y Secundaria* (pp. 9-18). Ediciones Universidad de Oviedo.

Gaudet, J., Gallant, F., & Bélanger, M. (2017). A bit of fit: Minimalist intervention in adolescents based on a physical activity tracker. *JMIR MHealth UHealth,5*(7), e92.https://doi.org/10.2196/mhealth.7647.

Guyatt, G., Oxman, A. D., Akl, E. A., Kunz, R., Vist, G., Brozek, J., Norris, S., Falck-Ytter, Y., Glasziou, P., DeBeer, H., Jaeschke, R., Rind, D., Meerpohl, J., Dahm, P., &Schünemann, H. J. (2011).

GRADEguidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *Journal Clinic of Epidemiology, 64*, 383-94. https://doi.org/10.1016/j.jclinepi.2010.04.026.

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers&Education, 80,* 152–161. https://doi.org/10.1016/j.compedu.2014.08.019

Hernández, A. J., Sánchez-Alcaraz, B. J., Alfonso-Asencio, M., & Hellín-Martínez, M. (2021). Gamificación en Educación Física. Revisión sistemática. *Trances, 14*(1), 1-20.

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., Welch, V. A. (2021). *Cochrane handbook for systematic reviews of interventions.* 6.2. Wiley.

Horne, P. J, Hardman, C. A., Lowe, C. F., Rowlands, A. V. (2009). Increasing children'sphysical activity: a peer modelling, rewards and pedometer-based intervention. *European Journal of Clinical Nutrition, 63*, 191-8. https://doi.org/10.1038/sj.ejcn.1602915.

Mayorga-Vega, D., Bocanegra-Parrilla, R., Ornelas, M., Viciana, J. (2016). Criterion-related validity of the distance- and time-based walk/run field testsfor estimating cardiorespiratory fitness: A systematic review and meta-analysis. *PLoS ONE, 11*, e0151671. https://doi.org/10.1371/journal.pone.0151671.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hofmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D.(2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, *372*(71), 1-9. https://doi.org/10.1136/bmj.n71.

Piñeiro Aboy, B. (2022). Los efectos de la Gamificación en el alumnado de Educación Física escolar. *International Multidisciplinary Journal CREA, 2,* 3-26.

Routen, A., Upton, D., Edwards, M., & Peters, D. (2014). The effect of pedometer stepgoal, feedback and self-monitoring interventions on accelerometer-measured physical activity in children. *Graduate Journal of Sport, Exercise & Physical Education Research, 2*, 37-53.

Werbach, K., & Hunter, D. (2012). *For the win: How game thinking can revolutionize your business.* Wharton Digital Press.