Abstract

Introduction: Poor diet, and the consumption of foods high in fat, sugar and salt are common causes of premature mortality and health conditions. Computerised response inhibition training has been proposed to devalue these foods by learning to inhibit motor responses, with effects found for both weight loss and snacking reduction. However, these interventions are repetitive by nature and suffer from a lack of adherence, leading some to propose gamification as a solution to increase engagement. The effect of gamification is unclear, however, with a lack of research investigating the effects of single game elements at improving adherence to interventions. The current study aims to investigate whether isolated common game elements improve adherence, engagement and effectiveness of computerised food response inhibition training compared to a standard non-gamified intervention. **Methods:** A sample of XXX participants were randomly assigned to either a standard non-gamified food response inhibition training, a training gamified with feedback elements, or a training gamified with social elements. Participants completed measures of snacking frequency and food evaluation before and after a 14-day training period, during which they were instructed to complete their assigned training. Training adherence and daily training motivation was recorded during the training period. **RESULTS: [The results will be described here]. Conclusions: The main conclusions drawn from the study will be stated here].**

TITLE: The effects of isolated game elements on adherence rates in food-based response inhibition training

Introduction

Poor diet is recognized as one of the leading causes of premature mortality (Roth et al., 2018), and though attempting to reduce calorific intake by dieting is common, it is often unsuccessful (Dohle et al., 2018). This may be due to prolonged calorie deficit increasing the reward value of high calorie food (Stice et al., 2019). Foods high in fat, sugar and salt (HFSS) generate strong reward responses that people susceptible to overeating can struggle to inhibit (Lawrence et al., 2022). Food response inhibition training (RIT) targets these automatic responses by training participants to inhibit motor responses to HFSS foods in computerised tasks (e.g., Go/No-Go tasks). Such training has been shown to reduce the consumption of the targeted unhealthy food in lab-based studies (Jones et al., 2016), facilitate short-term weight loss in field studies (Veling et al., 2014; Lawrence et al., 2015), and reduce the palatability of HFSS food items (Chen et al., 2016; Veling et al., 2017).

Food-specific RIT involves training participants to inhibit motor responses to unhealthy foods, pairing cues to withhold a motor response with HFSS stimuli, often in comparison with healthy food stimuli and non-food related controls (Jones et al., 2016). The mechanisms by which these interventions operate are still uncertain, though recent reviews propose that the learned motor response inhibition in response to HFSS foods conflicts with the reward responses elicited by the appetitive unhealthy food, which is resolved by a devaluation of the food stimulus (Chen et al., 2018; Lawrence et al., 2022; Quandt et al., 2019). In recent reviews, the effect of RIT on food devaluation has been demonstrated (Yang et al., 2022) and is supported by neuroimaging studies, which find reductions in activity in reward- and attention-related regions of the brain (Chen et al., 2018; Stice et al., 2017). Though there is promising evidence of the potential benefits of RIT, there are questions about adherence rates to computerised cognitive training delivered outside of the lab (Kelders et al., 2012; Lumsden et al., 2017), as inconsistent usage of RIT, or stopping altogether, can reduce the efficacy of training (Aulbach et al., 2021). Supporting this, Chen et al. (2019) found the effect of RIT on food choices to be reduced within weeks, suggesting the need for regular training.

RIT relies on the repetition of motor responses to similar stimuli hundreds of times over the course of an intervention, and engagement levels can waver over time. Engagement is not a well-defined term, but can be interpreted as both the experience of completing a task, and the participant’s behaviour when interacting with a task or intervention, such as how many sessions an individual completes or the timepoint at which they leave the study (Lumsden et al., 2017; Perski et al., 2017). Gamification - the process of adding game elements to tasks and systems without actually creating a game (Deterding et al., 2011) - presents a potential avenue to increase engagement with cognitive tasks and training. The rise in studies of gamified and game-like programs and tasks has resulted in several reviews to determine whether gamification can enhance intervention effectiveness, though the findings are mixed. Lumsden et al. (2017) and Vermeir et al. (2020) found tentative evidence that gamification can increase motivation and engagement with cognitive tasks, however whether this translates to increased adherence to an intervention schedule is unclear (Lumsden et al., 2017). Najberg et al. (2021) achieved protocol adherence rates of 95% with their food-based Go/NoGo and cued approach training, though there was no non-gamified control group, and the incentive for taking part may have been valuable enough to motivate participants to adhere to the protocol, with incentives known to improve recruitment and adherence rates (Abdelazeem et al., 2022). Aulbach et al. (2021) found a sharp decline in the number of participants using a gamified food RIT app (FoodT) over the first 5 days of use in an opportunistic study of real world usage with no incentive on offer. Furthermore, some studies have found that adding gamification elements may actually weaken training effects, which may be explained by game elements creating a distraction from the core stimuli (Forman et al., 2019). Careful consideration of how and when to add game-like elements is therefore important, but these elements are rarely examined in isolation (Seaborn & Fels, 2015; Vermeir et al., 2020), rendering their effects unclear. Game elements should be chosen from a theoretical perspective (Sailer et al., 2017; Vermeir et al., 2022), with self-determination theory (Deci & Ryan, 1985) commonly used to understand the potential role gamification elements play in increasing motivation and engagement in tasks.

Self-determination theory proposes that three psychological needs drive intrinsic motivation; that is motivation without a need for external reward. The need for competence refers to the desire to feel success when interacting with an environment; relatedness refers to a desire to belong to a group and feel coherent within a social structure; and autonomy refers to the desire to both be free to choose to perform an action, and feel performing that action is consistent with one's personal values (Sailer et al., 2017). Ryan et al. (2006) found these constructs predicted future video game-playing behaviour, though whether gamifying otherwise serious tasks affects these motivational constructs is an area needing more research (Seaborn & Fels, 2015). A survey of gamified work-related apps, such as those used for productivity and task management, found no effect of gamification on measures of autonomy, relatedness, or competence (Mitchell et al., 2021), though the apps were varied in purpose and nature of gamified elements. In a review of the computerised cognitive training literature, Vermeir et al. (2020) found the most common game elements to be those related to achievement and progression, such as point-based systems and feedback loops, and immersion elements such as a story or theme. These elements can be mapped to fulfilling psychological needs as set out by self-determination theory, with Sailer et al. (2017) finding specific elements (e.g. points and leader boards) were rated higher on corresponding psychological needs (e.g. competence). However, there remains a paucity of research investigating the effectiveness of single gamified elements (Mazarakis, 2021). Though recent research has investigated the effects of elements in isolation, as well as when different elements are combined in simple and single session tasks, further work is needed to determine the effects of isolated impact and whether they lead to changes in intrinsic motivation across more complex and longer interventions (Mazarakis & Brauer, 2022).

The current study, therefore, aims to examine the effects of isolated gamification elements on RIT engagement, adherence, and effectiveness in comparison to a well-established non-gamified intervention control. Our first gamified group adds social elements, by allowing participants to pick and join a team to contribute to, which is thought to increase motivation by fulfilling a need for social relatedness and providing a sense of relevance to their completion of the gamified task (Sailer et al., 2017). Our second gamified group adds feedback elements, thought to increase motivation by addressing a psychological need for competence according to self-determination theory (Sailer et al., 2017). There are four specific research questions each with associated hypotheses:

RQ1 – Does gamification lead to improved training adherence and training motivation compared to the standard, non-gamified version of food response inhibition training?

H1a – Gamified training groups will have a significantly greater number of completed sessions compared to the non-gamified training control group.

H1b – Gamified training groups will report higher levels of training motivation compared to the non-gamified training control group.

RQ2 – Does gamification improve training effects on food evaluations and snacking?

H2a – There will be a larger decrease in the liking ratings for unhealthy items in gamified groups compared to the non-gamified training control group.

H2b – There will be a larger increase in liking ratings for healthy foods in gamified groups compared to the non-gamified training control group.

H2c – The gamified training groups will display a greater reduction in unhealthy food item snacking in week following completion of the training compared to the control group.

RQ3 – Does training motivation and adherence mediate training response?

H3a – Pre- to Post-intervention differences in both unhealthy and healthy food item evaluations will be mediated by training adherence.

H3b - Pre- to Post-intervention differences in both unhealthy and healthy food item evaluations will be mediated by training motivation.

H3c – Pre- to Post-intervention differences in snacking frequency will be mediated by training adherence.

H3d - Pre- to Post-intervention differences in snacking frequency will be mediated by training motivation.

RQ4 – Is there equivalence between the gamification types for training adherence and motivation?

H4a – Training adherence rates will be equivalent between feedback and socially gamified training groups.

H4b – Training motivation will be equivalent between feedback and socially gamified training groups.

Given the lack of previous work on the effect of gamification on specific components of motivation, and potential equivalence of training effectiveness between single task gamification groups, we do not propose to test any hypotheses, however we do state our intention to explore the effects of gamification here to inform future research.

This project will also include measures of stress, personality, dietary behaviours, video game experience and inhibitory control for student dissertation projects, but these will not form part of any hypothesis tests or exploratory analysis we will report.

Methods

Transparency and Openness Statement

All data will be made openly available online via the University of Bath data repository archive, with materials and code available on the OSF site for this study: https://osf.io/jdk5f/?view_only=c87013dfff8e43aea702c8cc83b7d2e1. Upon receiving in principle acceptance, we will preregister this study on the Open Science Framework. In the sections below, we report all manipulations, measures, and exclusions. This study meets Level 6 of the PCI RR bias control (https://rr.peercommunityin.org/help/guide_for_authors).

Ethical Statement

This study is currently in the process of gaining ethical approval for each institution.

Design

This study will be a three-arm randomised controlled design, with intervention type as the 3-level grouping variable (non-gamified food inhibition training [control], achievement-related gamified inhibition training, social-oriented gamified inhibition training) and pre- and post-intervention as our repeated measures variable. Participants will be randomly allocated to groups using block randomisation (Moher et al., 2010), with a block size of 3, and blinded to the other training conditions in the study. Participants will be recruited via research participation schemes (e.g. SONA Systems Ltd), Prolific Academic (https://prolific.co/), and social media platforms (e.g. Twitter, Instagram).

Participants

The inclusion criteria for this study are as follows:

- Aged 18-65 with a Body Mass Index of 18.5 or above (suggesting a 'healthy' weight or above), consistent with previous research investigating food RIT training (e.g Lawrence et al., 2015).
- They report snacking on either crisps, chocolate, biscuits and/or cake (foods high in sugar, salt and fat) at least three times per week, as measured on an unhealthy snacking Food Frequency Questionnaire, consistent with previous work investigating the effect of computerised response inhibition training (e.g. Lawrence et al., 2015).
- Have access to a stable internet connection and a personal computer or laptop.

The exclusion criteria for this study are:

- A current or previous clinical diagnosis of an eating disorder or diabetes, or self-identifying as having either an eating disorder or diabetes.
- Currently attending a formal weight loss programme or using weight loss medication.

- Currently attempting to quit smoking, due to the changes in appetite and food craving during nicotine withdrawal (Kragel et al., 2019).

Sample Size Estimation

With our current resources, we estimate it is possible to recruit 80 participants per group, for a sample size of 240 in total. This would allow us to detect an effect size of *f* = 0.23 with 90% power. Given previous literature finding a large effect of gamification on task engagement, g = 0.72 (which we approximate to a Cohens f value of 0.36), with no evidence of publication bias (Vermeir et al., 2020), we believe this to be an appropriate target sample size that would yield informative results.

To measure our secondary hypotheses, an effect size of *f* = 0.24 was estimated for devaluation scores based on the previous work of the authors (from a *d* = 0.48, Lawrence et al., 2015). A G* Power (Faul et al., 2007) analysis to detect an interaction effect between 3 groups with 2 measurements, indicates that a total sample size of 60 is required to achieve 90% statistical power.

Finally, from a power analysis using the TOSTER R package (Lakens & Caldwell, 2021), we would be able to detect equivalence within the parameters *d* = -0.46 and *d* = 0.46 at 80% power with a sample size of 80 per group. We have been more lenient with our target power in this analysis to target relevant effect sizes which correspond to our previously stated effect size of interest (converting from *f* values of 0.23).

Our total target sample size is therefore set at 240.

The final sample comprised XXX participants (Mean Age = , SD = , % female, % White British), which had XX power to detect effects of XX. The study was ethically approved by each institution and all participants will provide informed consent.

Materials and Measures

Training Conditions

A non-gamified standard food response inhibition training will act as our active control group and is taken from Lawrence et al. (2015). Pictures of 18 food items, 9 being healthy (e.g. fruit or vegetables) and 9 being high-energy density foods (defined as being greater than 4 kcal/g such as cakes and chocolate), and 18 non-food items (clothes) will be presented on either the left of right of a screen. Stimuli will be presented for 1250ms, followed by a 1250ms inter-stimulus interval. Participants will press a key corresponding to the position of the stimuli on the screen ('c' for left and 'm' for right). Stimuli will be presented on a white screen with a frame at the border, which will turn bold on trials where the participant must inhibit their response ('no-go' trials, see Figure 1a). Healthy food items will always be paired with the 'go' instruction and unhealthy items will always be paired with the 'no-go' instruction, whilst non-food items will be associated with 'no-go' instructions in 50% of trials. Each of the 36 images will be presented once per block, with 6 blocks per training session.

A pseudo-socially gamified food response inhibition training will again be identical in nature to the standard response inhibition training, with the addition of a social element at the beginning and the end of each training session. Participants will select a team to join upon enrolment in the group: 'green', 'purple', 'yellow' or 'blue'. This will be their team for the duration of their training period, the border around the screen signalling a 'go' or 'no-go' trial changed to their selected team colour around the screen. At the beginning of each training session, they will be informed of their team position in a league table, as decided by a random number generator (see Figure 1b). This position will randomly increase by one, remain constant, or decrease by one at the end of each session. Socially oriented gamified elements are among the least researched elements (Vermeir et al., 2020).

A feedback gamified food response inhibition training will be identical in nature to the training protocol described previously, with the addition of points awarded or deducted based on their performance in each block. 10 points will be awarded for correct 'go' responses, commission errors will result in a 5 point deduction, with visual confirmation of each provided after each trial (see Figure 1c). At the end of each training session, participants will be awarded a badge, decorated with a number of stars corresponding to the number of training sessions in a row they have completed. This manipulation is chosen as it is a frequently included gamification element (e.g. Vermeir et al., 2020, Vermeir et al., 2022).

INSERT FIGURE 1a.

INSERT FIGURE 1b.

INSERT FIGURE 1c.

Confirmatory Outcome Variables

Food liking and value judgement confidence will be measured with a food evaluation task, in accordance with previous research (Lawrence et al., 2015). 27 pictures of food will be presented, 9 healthy food and 9 high-energy density food images included in the inhibition training, with the other 9 novel food images. Participants will be asked to rate how much they like each food image on a 100mm visual analogue scale, anchored at the extremes with "not at all" (0) and "very much" (100), and then how confident they are in the evaluation they just gave. Participants will select their rating on the scale with their mouse cursor, which will be set to the midpoint (a rating of 50) at the start of each trial. Both liking and confidence ratings will be collected pre- and post-intervention.

Snacking frequency will be measured with a Snacking Food Frequency Questionnaire (FFQ; Churchill & Jessop, 2011). This presents as an 8-item, 7-point Likert-type scale, asking about how frequently different foods (namely crisps, cakes, chocolate, and biscuits) were consumed over the previous 2 weeks. Answers can range from 1 (“not at all”) to 7 (“More than 4 times per week”). Responses are summed to form a total score, with higher scores indicating greater snacking.

Adherence will be measured as the number of training sessions out of 10, completed by a participant with an accuracy of >80% on both go and no-go trials.

Training motivation and enjoyment will be measured by a questionnaire given at the end of each training session. Scores for each questionnaire will be summed to give a daily score, with overall training motivation approximated as the mean of all daily scores. The questionnaire has been adapted from Lumsden et al. (2017) and includes the following items: 1. How enjoyable did you find that? 2. How frustrating did you find that? 3. How mentally stimulating did you find that? 4. How repetitive did you find that? 5. How willing would you be to do that again tomorrow? Participants will record their answer to each item on a visual analogue scale 100mm long with “not at all” at one end, to “very” at the other with no subdivisions. After reverse scoring questions 2 and 4, mean item scores will be calculated for each participant, with higher values indicating greater subjective enjoyment.

<u>Exploratory Outcome Variables</u>

Intrinsic motivation will be assessed by the Intrinsic Motivation Inventory (IMI; Deci & Ryan, 2005), to refer specifically to the training task (see Appendix A). This is a 19-item questionnaire with answers ranging from 1 (“Not all”) to 7 (“Very true”) with questions measuring relatedness (e.g. “I felt connected with the others taking part in this study”), autonomy (e.g. “I believe I had some choice about completing each training session”), and competence (“I think I did pretty well at this training, compared to other participants”) After reverse scoring relevant items, scores will be summed to created total scores for each subscale, with higher scores indicating greater sense of relatedness, autonomy and competence.

Lifestyle factors will be assessed with a 100mm visual analogue scale as before, set at the midpoint (a value of 50), anchored at either end with “No” (0) and “Yes” (100). questions asking about a participant’s sleep quality from the previous night (“Did you get enough sleep last night?”), whether they are currently hungry ("Are you currently hungry?") and whether they are hungover ("Are you currently hungover") prior to training or testing sessions.

Attention checks

To protect against careless responding (Jones et al., 2022) participants will be asked the multiple choice question "What planet do you live on? Earth, Mars, Mercury, Saturn" after the food evaluation task (taken from Pennington et al., 2023). Analyses will be conducted with and without those who fail this attention check, with differences reported.

Procedure

This study will be designed and hosted on the Gorilla online platform (Gorilla.sc).

Potential participants will first complete a screening stage consisting of completing the Food Frequency Questionnaire (Churchill & Jessop, 2011). They will also be asked their height and weight, and this will be converted to yield their BMI (kg/m2). Data collection will occur over two testing sessions: baseline and post-intervention. Heights below 122cm and above 213cm (48 and 84 inches), or weights below 34kg and above 227kg (75 and 500 pounds) will be deemed implausible, and therefore not included in any descriptive statistics (Das et al., 2005). At baseline, we will record participants' self-reported weight, food evaluations and snacking frequency, alongside measures of stress, personality, dietary behaviours and inhibitory control for student dissertation projects. Between the two testing sessions, participants are asked to complete 10 sessions of an online training intervention over 14-days. 7 days after this intervention period, participants will be asked to complete the food evaluation measure. Finally, participants will compete all measures again post-intervention. Participants will be paid £6 per hour, in line with the guidelines of Prolific Academic and the Lead institution.

Planned Statistical Analyses

Data Exclusion Criteria

Outliers on outcome measures will be retained for the analysis and participants who complete less than 10 training sessions will be retained in order to achieve an unbiased effect estimate of the intervention (McCoy, 2017). Participants who average below 80% accuracy on training sessions or whose accuracy score is greater than 2 standard deviations higher than the group mean will be excluded from the analysis.

Missing data will be assessed for randomness using Little's test (Little, 1986), and a missing dummy variable will be created and tested for associations with group type. Complete case analyses will be carried out should less than 5% of the sample data be missing or the data is assessed as missing completely at random. Should either test not be met, multiple imputation using the r package 'mice' (Buuren et al., 2023) will be carried out to replace missing data points, with 5 imputations carried out based on all available data.

Planned Analyses

All analyses will be carried out in R Studio, using the packages dplyr, mice, mediation, tidyr, lmer, TOSTER, naniar, apatables and papaja. Significance will be accepted when $p < .05$ for all tests.

RQ1 – Does gamification lead to improved training adherence and motivation compared to non-gamified training?

All hypotheses relating to this question will be tested with a one-way factorial ANOVA with training group as the factor, and mean motivation score and number of training sessions completed as our dependent variables respectively. Should this result be significant, we will follow this up with independent-samples t-tests.

RQ2 – Does gamification improve training effects for food evaluations and snacking?

All hypotheses addressing this research question will be tested with 3 (Control, Social Gamification, Reward Gamification) x 2 (Time, Pre – Post) mixed design Analysis of Variance tests (ANOVAs). Should the interaction term, and main effects be significant we will follow this up with analysis of simple main effects to investigate the direction of the effect found.

RQ3 – Does motivation and adherence mediate training response?

Mediation analyses will be carried out following using a causal steps approach, as suggested by Hayes and Rockwood (2017), for each hypothesis separately. Significance of the mediation model will be determined using the bootstrap method, based on 5000 bootstrap samples (consistent with the recommended number from Hayes, 2018). Mediation coefficients will be established with separate regressions: first with intervention group entered as a predictor (dummy coded with the control set as the reference category) and change in food item evaluations, and in unhealthy food snacking frequency, as the outcome. Secondly the direct effect of intervention group on change in training engagement or motivation will be established, followed by establishing the indirect effect with both intervention group and change in the mediator as our predictor variables and change in food evaluation score, and snacking frequency, as our outcome variable. For all models, baseline scores of our variables will be entered as covariates and to establish the significance and confidence intervals, the R package '*mediation*' will be used (Tingley et al., 2014).

RQ4 – Is there equivalence between the gamification types for training adherence and motivation?

All hypotheses for this research question will be tested with two one-sided t-tests, using the R package *TOSTER* (Lakens & Caldwell, 2021). Equivalence will be assumed should 95% confidence intervals fall between these parameters.

Exploratory Analysis

Whilst we are not going to formally test a specific hypothesis, we will also investigate whether gamification results in fewer outliers (scores greater than 2 standard deviations from the group mean) during training performance.

Secondly, we will explore whether there are differences between the training groups on measures of perceived competence, autonomy, and relatedness, in line with Sailer et al.'s (2017) categorisation of gamification elements with self-determination theory.

Thirdly, we will explore whether gender moderates the relationship between gender and adherence rates based on previous findings that suggest gender moderates training efficacy in a largely theme-based gamification task (Forman et al., 2021).

Finally, we will explore relationships between change in food item evaluation confidence, food liking ratings and training adherence.

References

Abdelazeem, B., Abbas, K. S., Amin, M. A., El-Shahat, N. A., Malik, B., Kalantary, A., & Eltobgy, M. (2022). The effectiveness of incentives for research participation: A systematic review and meta-analysis of randomized controlled trials. *PLOS ONE*, *17*(4), e0267534. https://doi.org/10.1371/journal.pone.0267534

Aulbach, M. B., Knittle, K., van Beurden, S. B., Haukkala, A., & Lawrence, N. S. (2021). App-based food Go/No-Go training: User engagement and dietary intake in an opportunistic observational study. *Appetite*, *165*, 105315. https://doi.org/10.1016/j.appet.2021.105315

Buuren, S. van, Groothuis-Oudshoorn, K., Vink, G., Schouten, R., Robitzsch, A., Rockenschaub, P., Doove, L., Jolani, S., Moreno-Betancur, M., White, I., Gaffert, P., Meinfelder, F., Gray, B., Arel-Bundock, V., Cai, M., Volker, T., Costantini, E., Lissa, C. van, & Oberman, H. (2023). *mice: Multivariate Imputation by Chained Equations* (3.16.0). https://cran.r-project.org/web/packages/mice/index.html

Chen, Z., Holland, R. W., Quandt, J., Dijksterhuis, A., & Veling, H. (2019). When mere action versus inaction leads to robust preference change. *Journal of Personality and Social Psychology*, *117*, 721–740. https://doi.org/10.1037/pspa0000158

Chen, Z., Veling, H., Dijksterhuis, A., & Holland, R. W. (2016). How does not responding to appetitive stimuli cause devaluation: Evaluative conditioning or response inhibition? *Journal of Experimental Psychology: General*, *145*, 1687–1701. https://doi.org/10.1037/xge0000236

Chen, Z., Veling, H., Dijksterhuis, A., & Holland, R. W. (2018). Do impulsive individuals benefit more from food go/no-go training? Testing the role of inhibition capacity in the no-go devaluation effect. *Appetite*, *124*, 99–110. https://doi.org/10.1016/j.appet.2017.04.024

Das, S. R., Kinsinger, L. S., Yancy, W. S., Wang, A., Ciesco, E., Burdick, M., & Yevich, S.

J. (2005). Obesity prevalence among veterans at Veterans Affairs medical facilities. *American Journal of Preventive Medicine*, *28*(3), 291–294. https://doi.org/10.1016/j.amepre.2004.12.007

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer US. https://doi.org/10.1007/978-1-4899-2271-7

Dohle, S., Diel, K., & Hofmann, W. (2018). Executive functions and the self-regulation of eating behavior: A review. *Appetite*, *124*, 4–9. https://doi.org/10.1016/j.appet.2017.05.041

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Forman, E. M., Manasse, S. M., Dallal, D. H., Crochiere, R. J., Berry, M. P., Butryn, M. L., & Juarascio, A. S. (2021). Gender differences in the effect of gamification on weight loss during a daily, neurocognitive training program. *Translational Behavioral Medicine*, *11*(4), 1015–1022. https://doi.org/10.1093/tbm/ibaa091

Forman, E. M., Manasse, S. M., Dallal, D. H., Crochiere, Rebecca. J., Loyka, C. M., Butryn, M. L., Juarascio, A. S., & Houben, K. (2019). Computerized neurocognitive training for improving dietary health and facilitating weight loss. *Journal of Behavioral Medicine*, *42*(6), 1029–1040. https://doi.org/10.1007/s10865-019-00024-5

Froud, R., Eldridge, S., Lall, R., & Underwood, M. (2009). Estimating the number needed to treat from continuous outcomes in randomised controlled trials: Methodological challenges and worked example using data from the UK Back Pain Exercise and Manipulation (BEAM) trial. *BMC Medical Research Methodology*, *9*(1), 35. https://doi.org/10.1186/1471-2288-9-35

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process* *analysis: A regression-based approach*. Guilford publications.

Hayes, A. F., & Rockwood, N. J. (2017). Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation. *Behaviour Research and Therapy*, *98*, 39–57. https://doi.org/10.1016/j.brat.2016.11.001

Johnson, D., Deterding, S., Kuhn, K.-A., Staneva, A., Stoyanov, S., & Hides, L. (2016). Gamification for health and wellbeing: A systematic review of the literature. *Internet Interventions*, *6*, 89–106. https://doi.org/10.1016/j.invent.2016.10.002

Jones, A., Di Lemma, L. C. G., Robinson, E., Christiansen, P., Nolan, S., Tudur-Smith, C., & Field, M. (2016). Inhibitory control training for appetitive behaviour change: A meta-analytic investigation of mechanisms of action and moderators of effectiveness. *Appetite*, *97*, 16–28. https://doi.org/10.1016/j.appet.2015.11.013

Jones, A., Hardman, C. A., Devlin, N., Pennington, C. R., & Robinson, E. (2022). Weight-based discrimination in financial reward and punishment decision making: Causal evidence using a novel experimental paradigm. *International Journal of Obesity*, *46*(7), Article 7. https://doi.org/10.1038/s41366-022-01109-z

Kelders, S. M., Kok, R. N., Ossebaard, H. C., & Gemert-Pijnen, J. E. V. (2012). Persuasive System Design Does Matter: A Systematic Review of Adherence to Web-Based Interventions. *Journal of Medical Internet Research*, *14*(6), e2104. https://doi.org/10.2196/jmir.2104

Kragel, E. A., Sweitzer, M. M., & Davis, J. M. (2019). The Effect of Brief Mindfulness Training on Brain Reactivity to Food Cues During Nicotine Withdrawal: A Pilot Functional Imaging Study. *Mindfulness*, *10*(11), 2272–2276. https://doi.org/10.1007/s12671-019-01201-y

Lakens, D., & Caldwell, A. (2021). *TOSTER: Two One-Sided Tests (TOST) Equivalence*

Testing (0.7.1). https://cran.r-project.org/web/packages/TOSTER/index.html

Lawrence, N. S., O'Sullivan, J., Parslow, D., Javaid, M., Adams, R. C., Chambers, C. D., Kos, K., & Verbruggen, F. (2015). Training response inhibition to food is associated with weight loss and reduced energy intake. *Appetite, 95*, 17–28. https://doi.org/10.1016/j.appet.2015.06.009

Lawrence, N. S., Porter, L., & Staiger, P. K. (2022). The 'Go's and the 'No-Go's of response-inhibition training to food: Lessons learned from trials. *Current Opinion in Behavioral Sciences*, *48*, 101229. https://doi.org/10.1016/j.cobeha.2022.101229

Little, R. J. A. (1988). A Test of Missing Completely at Random for Multivariate Data with Missing Values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Lumsden, J., Skinner, A., Coyle, D., Lawrence, N., & Munafo, M. (2017). Attrition from Web-Based Cognitive Testing: A Repeated Measures Comparison of Gamification Techniques. *Journal of Medical Internet Research*, *19*(11), e8473. https://doi.org/10.2196/jmir.8473

Mazarakis, A. (2021). Gamification Reloaded. *Current and Future Trends in Gamification Science*, *20*(3), 279–294. https://doi.org/10.1515/icom-2021-0025

Mazarakis, A., & Bräuer, P. (2023). Gamification is Working, but Which One Exactly? Results from an Experiment with Four Game Design Elements. *International Journal of Human–Computer Interaction*, *39*(3), 612–627. https://doi.org/10.1080/10447318.2022.2041909

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., Elbourne, D., Egger, M., & Altman, D. G. (2010). CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. *BMJ*, *340,* c869. https://doi.org/10.1136/bmj.c869

Morey, R. D., Rouder, J. N., Jamil, T., Urbanek, S., Forner, K., & Ly, A. (2022). *BayesFactor: Computation of Bayes Factors for Common Designs* (0.9.12-4.4). https://cran.r-project.org/web/packages/BayesFactor/index.html

Najberg, H., Rigamonti, M., Mouthon, M., & Spierer, L. (2021). Modifying food items valuation and weight with gamified executive control training. *Royal Society Open Science*, *8*(5), 191288. https://doi.org/10.1098/rsos.191288

Pennington, C. R., Monk, R. L., Heim, D., Rose, A. K., Gough, T., Clarke, R., Knibb, G., Patel, R., Rai, P., Ravat, H., Ali, R., Anastasiou, G., Asgari, F., Bate, E., Bourke, T., Boyles, J., Campbell, A., Fowler, N., Hester, S., … Jones, A. (2023). The labels and models used to describe problematic substance use impact discrete elements of stigma: A registered report. *Psychology of Addictive Behaviors*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/adb0000919

Perski, O., Blandford, A., West, R., & Michie, S. (2017). Conceptualising engagement with digital behaviour change interventions: A systematic review using principles from critical interpretive synthesis. *Translational Behavioral Medicine*, *7*(2), 254–267. https://doi.org/10.1007/s13142-016-0453-1

Porter, L., Gillison, F. B., Wright, K. A., Verbruggen, F., & Lawrence, N. S. (n.d.). *Exploring Strategies to Optimise the Impact of Food-Specific Inhibition Training on Children's Food Choices*.

Quandt, J., Holland, R. W., Chen, Z., & Veling, H. (2019). The role of attention in explaining the no-go devaluation effect: Effects on appetitive food items. *Journal of Experimental Psychology: Human Perception and Performance*, *45*, 1119–1133. https://doi.org/10.1037/xhp0000659

Roth, G. A., Mensah, G. A., Johnson, C. O., Addolorato, G., Ammirati, E., Baddour, L. M., Barengo, N. C., Beaton, A. Z., Benjamin, E. J., Benziger, C. P., Bonny, A., Brauer, M., Brodmann, M., Cahill, T. J., Carapetis, J., Catapano, A. L., Chugh, S. S., Cooper, L. T., Coresh, J., … Fuster, V. (2020). Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019: Update From the GBD 2019 Study. *Journal of the American College of Cardiology*, *76*(25), 2982–3021. https://doi.org/10.1016/j.jacc.2020.11.010

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, *69*, 371–380. https://doi.org/10.1016/j.chb.2016.12.033

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *74*, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006

Stice, E., Yokum, S., Nelson, T. D., Berkman, E., Veling, H., & Lawrence, N. (2022). Efficacy of a combined food-response inhibition and attention training for weight loss. *Current Opinion in Behavioral Sciences*, *46*, 101168. https://doi.org/10.1016/j.cobeha.2022.101168

Stice, E., Yokum, S., & Waters, A. (2015). Dissonance-Based Eating Disorder Prevention Program Reduces Reward Region Response to Thin Models; How Actions Shape Valuation. *PLOS ONE*, *10*(12), e0144530. https://doi.org/10.1371/journal.pone.0144530

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., Imai, K., Trinh, M., & Wong, W. (2019). *mediation: Causal Mediation Analysis* (4.5.0). https://cran.r-project.org/web/packages/mediation/index.html

Tzavella, L., Lawrence, N. S., Button, K. S., Hart, E. A., Holmes, N. M., Houghton, K., Badkar, N., Macey, E., Braggins, A.-J., Murray, F. C., Chambers, C. D., & Adams, R.

C. (2021). Effects of go/no-go training on food-related action tendencies, liking and choice. *Royal Society Open Science*, *8*(8), 210666. https://doi.org/10.1098/rsos.210666

Veling, H., Lawrence, N. S., Chen, Z., van Koningsbruggen, G. M., & Holland, R. W. (2017). What Is Trained During Food Go/No-Go Training? A Review Focusing on Mechanisms and a Research Agenda. *Current Addiction Reports*, *4*(1), 35–41. https://doi.org/10.1007/s40429-017-0131-5

Veling, H., van Koningsbruggen, G. M., Aarts, H., & Stroebe, W. (2014). Targeting impulsive processes of eating behavior via the internet. Effects on body weight. *Appetite*, *78*, 102–109. https://doi.org/10.1016/j.appet.2014.03.014

Vermeir, J. F., White, M. J., Johnson, D., Crombez, G., & Van Ryckeghem, D. M. L. (2020). The Effects of Gamification on Computerized Cognitive Training: Systematic Review and Meta-Analysis. *JMIR Serious Games*, *8*(3), e18644. https://doi.org/10.2196/18644

Vermeir, J. F., White, M. J., Johnson, D., Crombez, G., & Van Ryckeghem, D. M. L. (2022). Gamified Web-Delivered Attentional Bias Modification Training for Adults With Chronic Pain: Protocol for a Randomized, Double-blind, Placebo-Controlled Trial. *JMIR Res Protoc*, *11*(1), e32359. https://doi.org/10.2196/32359

Yang, Y., Qi, L., Morys, F., Wu, Q., & Chen, H. (2022). Food-Specific Inhibition Training for Food Devaluation: A Meta-Analysis. *Nutrients*, *14*(7).

Figures

Figure 1a. Example Standard Non-Gamified Food RIT Control 'Go' Stimulus.

Figure 1b. Example Socially Gamified Food RIT League Table screen.

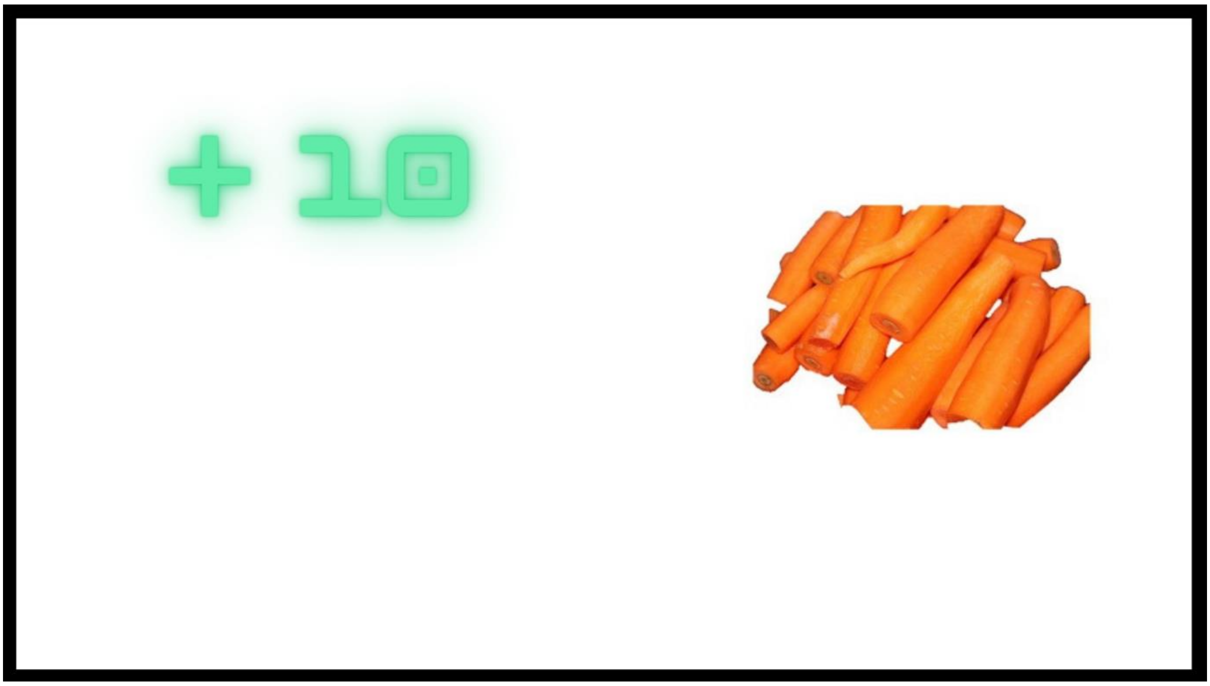


Figure 1c. Example Feedback Gamified Food RIT 'No-Go' correct response points feedback

Appendix A - Scales

Intrinsic Motivation Inventory for RIT

Relatedness

- I felt really distant to the others taking part in this study. (R)
- I really doubt that I have much in common with others taking part in this study. (R)
- I felt connected with the others taking part in this study.
- I felt like the others taking part in this study and I were working towards a common goal.
- It is likely that the other people taking part and I could become friends if we interacted a lot.
- I felt close to my team. *(Social condition only, excluded from total scores)*

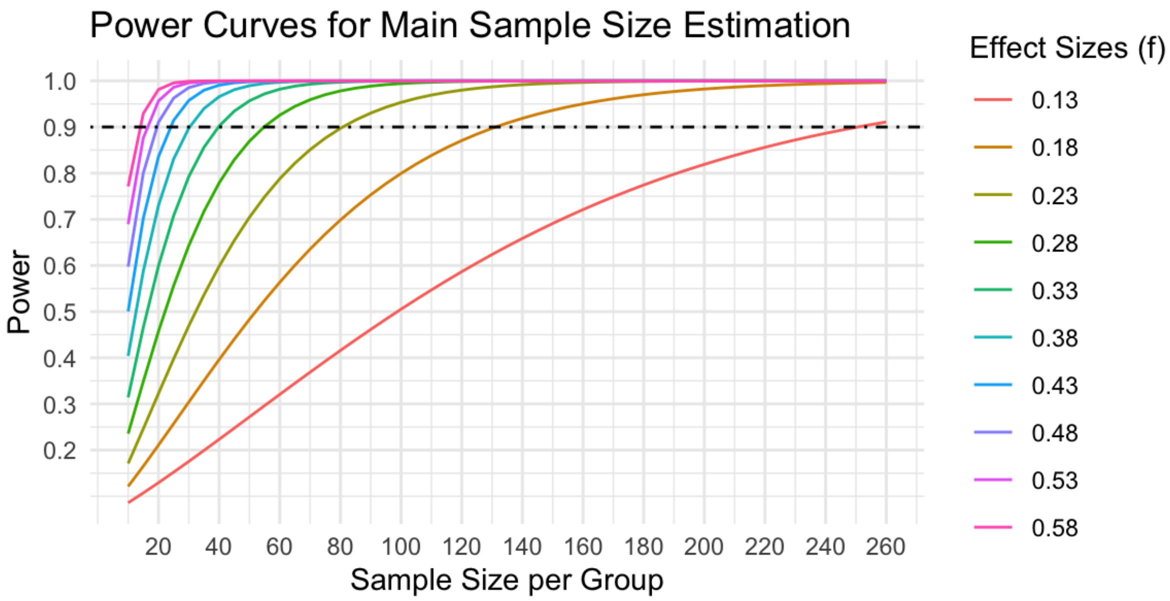
Perceived Competence

- I think I am pretty good at this training.
- I think I did pretty well at this training, compared to other participants.
- After working at this training for a while, I felt pretty competent.
- I am satisfied with my performance over the course of the training period.
- I was pretty skilled at this training task.
- This was a task that I couldn’t do very well. (R)

Perceived Choice

- I believe I had some choice about completing each training session.
- I felt like it was not my own choice to do a training session. (R)
- I didn’t really have a choice about taking part in the training. (R)
- I felt like I had to do this. (R)
- I did this because I had no choice. (R)
- I did each training session because I wanted to.
- I did each training session because I had to. (R)

Appendix B – Power Curves for Sample Size Estimation



Question	Hypothesis	Sampling plan	Analysis Plan	Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis	Interpretation given different outcomes	Theory that could be shown wrong by the outcomes
Does gamification lead to improved training adherence and training motivation compared to the standard, non-gamified version of food response inhibition training?	H1a – Gamified training groups will have a significantly greater number of completed sessions compared to the non-gamified training control group (primary hypothesis). H1b – Gamified training groups will report higher levels of training motivation compared to the non-gamified training control group.	Based on an effect size of f = 0.23 we propose to recruit 80 participants per group, 240 in total, to detect between group differences on adherence and engagement rates.	All hypotheses relating to this question will be tested with a one-way factorial ANOVA with training group as the factor, and mean motivation score and number of training sessions completed as our dependent variables respectively.	All hypotheses will be tested using frequentist methods with significance will be accepted at the $p < 0.05$ level.	Should these hypotheses be confirmed, we will interpret this as evidence for simple gamification improving adherence and/or enjoyment of food RIT Should we fail to reject the null, we will interpret this as a lack of evidence for simple gamification improving adherence and enjoyment of food RIT.	Theories that propose gamification renders tasks more motivating and engaging will be challenged by a null finding.
Does gamification improve training effects on food	H2a – There will be a larger decrease in the liking ratings for unhealthy items	Based on an effect size of f = 0.24, we propose to recruit 24 participants per	All hypotheses addressing this research question will be tested with		Should these hypotheses be confirmed, we will interpret this as	Should we reject the null, this would contrast with previous research

evaluations and snacking?	in gamified groups compared to the non-gamified training control group. H2b – There will be a larger increase in liking ratings for healthy foods in gamified groups compared to the non-gamified training control group. H2c – Gamified training groups will display a greater reduction in unhealthy food item snacking in week following completion of the training compared to the control group.	group, for a total of 72 to detect group x time (pre-post) interaction effects.	3 (Control, Social Gamification, Reward Gamification) x 2 (Time, Pre – Post) mixed design Analysis of Variance tests (ANOVAs)		evidence for simple gamification improving the effectiveness of food RIT Should we fail to reject the null, we will interpret this as a lack of evidence for simple gamification improving adherence and enjoyment of food RIT.	finding negative effects of gamification on the effectiveness of cognitive training, namely food RIT (e.g. Forman et al., 2019)
RQ3 – Does training motivation and adherence mediate training response?	H3a – Pre- to Post-intervention differences in both unhealthy and healthy food item	No specific sampling plan has been carried out for this analysis.	Mediation analyses will be carried out following using a causal steps	Significance of the mediation model will be determined using the bootstrap	Should we reject the null, this would be interpreted as evidence that	A null result here may challenge theories that gamification may

	evaluations will be mediated by training adherence. H3b - Pre- to Post-intervention differences in both unhealthy and healthy food item evaluations will be mediated by training motivation.		approach, as suggested by Hayes and Rockwood (2017), for each hypothesis separately.	method, based on 5000 bootstrap samples (consistent with the recommended number from Hayes, 2018).	gamification improves training effectiveness by increasing user engagement, however this will be limited due to sample size constraints. Should we fail to reject the null, this shall be interpreted as a lack of evidence for gamifying food RITs with the aim of increasing effectiveness by increasing engagement.	improve training effects by increasing engagement and/or adherence.
RQ4 – Is there equivalence between the gamification types for training adherence and motivation?	H4a – Training adherence rates will be equivalent between feedback and socially gamified training groups.	For our main frequentist tests, we were powering to detect an effect size of f = 0.23 (an equivalent Cohen's d value of 0.46). Though this	All hypotheses for this research question will be tested with two one-sided t-tests, using the TOSTER R package.	Equivalence will be accepted should both the equivalence test result return as significant. Should the null hypothesis test return as	Should we reject the null, we will determine there is no meaningful difference between gamification types on either	No specific theories will be challenged by results here, though future avenues for research will be suggested should

	H4b – Training motivation will be equivalent between feedback and socially gamified training groups.	hypothesis was not part of our formal sample size calculation, using the TOSTER power function, we calculate that 80 participants per group is sufficient to detect effects at this size with 80% power.		significant and the equivalence test return as not significant, this will not be interpreted as equivalence between the two groups.	adherence or engagement levels. Should we fail to reject the null, we will interpret this as evidence that gamified elements affect adherence and/or engagement differently.	we fail to reject the null in this case.