TILBURG UNIVERSITY

Tilburg University

Commitment Lotteries

van der Swaluw, Koen

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Link to publication in Tilburg University Research Portal

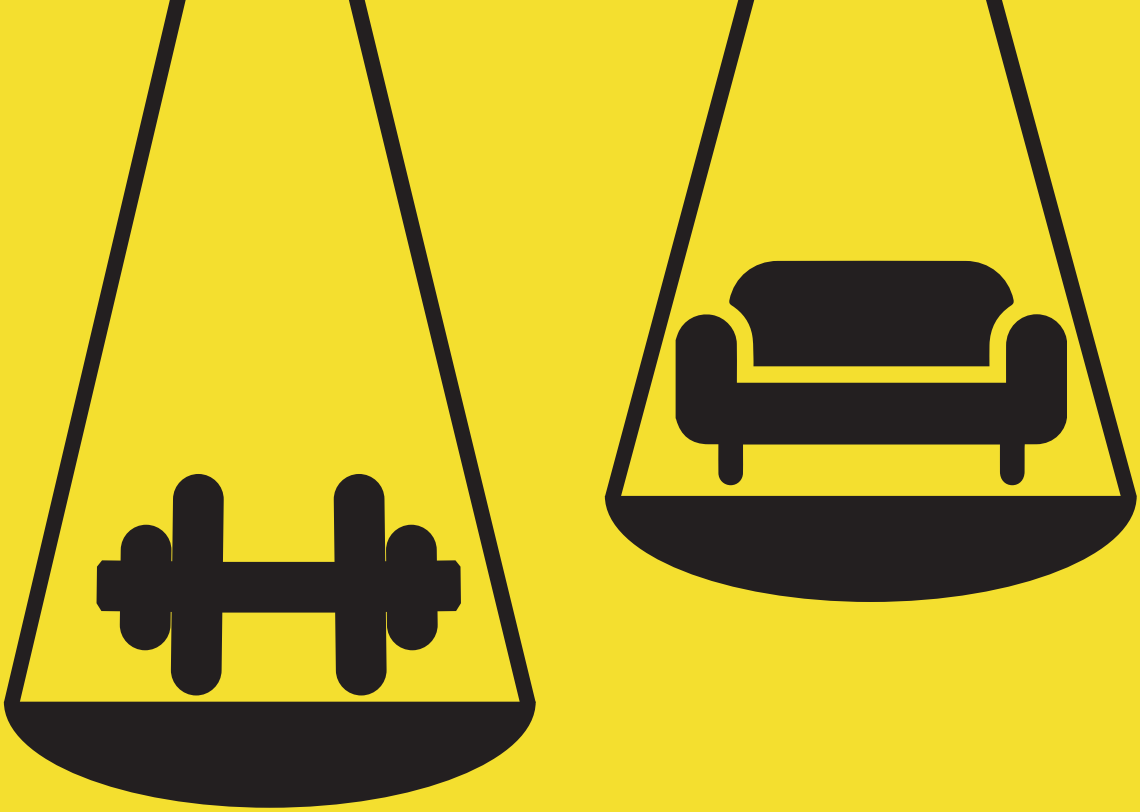
Citation for published version (APA):
van der Swaluw, K. (2018). *Commitment Lotteries: Overcoming procrastination of lifestyle improvement with regret aversion*.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 12. May. 2021



COMMITMENT LOTTERIES

Overcoming procrastination of lifestyle improvement with regret aversion

Koen van der Swaluw

COMMITMENT LOTTERIES

Overcoming procrastination of lifestyle improvement with regret aversion

Koen van der Swaluw

Colofon

Commitment lotteries. Overcoming procrastination of lifestyle improvement with regret aversion Koen van der Swaluw
ISBN/EAN: 978-94-028-1269-5

Copyright © 2018 Koen van der Swaluw
All rights reserved. No part of this thesis may be reproduced, stored or transmitted in any way or by any means without the prior permission of the author, or when applicable, of the publishers of the scientific papers.

The research reported in this dissertation was financially supported by the National lnstitute for Public Health and the Environment (RlVM, SPR program Health Economics) and Tilburg University.

Layout and design by Douwe Oppewal
Printed by lpskamp Printing

Commitment Lotteries

Overcoming procrastination of lifestyle improvement with regret aversion

Proefschrift

ter verkrijging van de graad van doctor
aan Tilburg University
op gezag van de rector magnificus, prof. dr. E.H.L. Aarts,
in het openbaar te verdedigen
ten overstaan van een door het college voor promoties aangewezen commissie
in de aula van de Universiteit op vrijdag
14 december 2018 om 10.00 uur

door

Koen van der Swaluw,
geboren te Wageningen

Promotores
Prof. dr. J. J. Polder
Prof. dr. H.M. Prast

Copromotores
Dr. M.S. Lambooij
Dr. J.J.P. Mathijssen

Promotiecommissie
Prof. dr. W. van Dijk
Prof. dr. A.J. Dijksterhuis
Prof. dr. N.J.A. van Exel
Prof. dr. L.A.M. van de Goor
Prof. dr. E.C.M. van der Heijden
Prof. dr. E.M. Sent

CONTENTS

CHAPTER I

General introduction

Disease prevention through improved lifestyle behavior is receiving increasing attention from policymakers, medical professionals and science (Hulsegge et al., 2016; Rijksoverheid, 2018; Van Winkelhof, Pijl, & Vliegenthart, 2018). Likewise, the fields of health economics, psychology and behavioral economics progressively understand the determinants of health related behavior (Bickel, Moody, & Higgins, 2016; Kooreman & Prast, 2010). Knowledge of the latter is essential for public health policy and practice. Operating at the crossroads of health economics, psychology and behavioral economics, this PhD thesis focuses on supporting individual lifestyle decisions.

Behavior as a determinant of health

From the late nineteenth century and throughout the first half of the twentieth century, population health developed impressively in Western nations (OECD, 2015). Developments in the areas of hygiene, housing and nutrition all substantially improved population health (Van der Lucht & Polder, 2010; Van Zon, 1990). From the Second World War, vaccinations, antibiotics and medicinal prevention boosted Dutch life expectancy up to 81.5 years in 2016 (Aminov, 2010; CBS, 2018). Many diseases that once threatened human health and well-being became preventable or curable through societal and medical developments.

In the twentieth century, the primary causes of death shifted from infectious diseases to so-called 'civilization diseases' (Van der Lucht & Polder, 2010). Today, seventy percent of deaths worldwide result from non-communicable diseases (NCDs), such as cardiovascular disease, type 2 diabetes, and multiple types of cancer (Forouzanfar et al., 2016). Hence, the United Nations (UN) resolution on the prevention and control of NCDs stresses the urgency of "multilateral efforts at the highest political level to address the rising prevalence, morbidity, and mortality" of NCDs globally (UN, 2010, p3.). In their 2030 Agenda for Sustainable Development the UN aims to "By 2030, reduce by one-third premature mortality from NCDs through prevention and treatment" (UN, 2015).

The World Health Organization (WHO) has established that most NCDs result from an unhealthy diet, insufficient physical activity, smoking and the harmful consumption of alcohol (Forouzanfar et al., 2016). Unfortunately, levels of physical activity have been falling (Lee et al., 2012), and global rates of obesity have more than doubled since 1980 (WHO, 2015). In Europe, two in every three citizens do not meet recommended levels of physical activity and 62% of Europeans are overweight or obese (EC, 2014).

A persistent level of premature mortality

Over the last two decades, there have been many European (governmental) initiatives to prevent amendable mortality, which is defined as premature deaths that could have been avoided through timely and effective health care. The Netherlands is one the best scoring countries in Europe in this aspect, which, according to the OECD points to steady improvements in the

access to and quality of health care (OECD, 2017). However, when it comes to deaths from preventable causes, the OECD notes that a persistent level of premature mortality remains from causes that could be prevented through improved lifestyle behavior (OECD, 2017).

In the Netherlands, 18.5% of healthy life years lost (DALY's) is linked to behavior: after smoking, unhealthy nutrition and physical inactivity are the two leading contributors (RIVM, 2018). Half of the Dutch population is overweight (Body Mass Index ≥ 25), 53% does not meet recommended levels of physical activity and about 43% exercises less than once per week (CBS & RIVM, 2017). In accordance, leading scholars in the field of social science have listed "How can we help people take care of their health?" as the most pressing question that social scientists should tackle nowadays (Giles, 2011).

People remain inactive

To a large extent, people know that their decisions are related to their health and often intend to improve their lifestyle (Kooreman & Prast, 2010). In the Netherlands, the most-mentioned resolution for 2018 was 'to exercise more' and the third most mentioned resolution was to lose weight (ING, 2017). Accordingly, books on diet and exercise remain consistently among the best sold books in the Netherlands (CPNB, 2018). Between 2010 and 2017, the Dutch bought €63 million worth of diet books (KVB, 2017). However, despite knowledge, intentions and the apparent willingness to pay for lifestyle improvements, people typically exercise much less than they initially intended (Acland & Levy, 2015; Carrera, Royer, Stehr, & Sydnor, 2018) and most weight loss attempts fail (Elfhag & Rössner, 2005). Similarly, many preventive measures by governments and organizations face the stubborn reality of human behavior and accomplish considerably less than anticipated (Van den Berg & Schoemaker, 2010).

Conventional theories cannot fully explain health damaging choices

To some degree, this can be explained by conventional assumptions about rational, maximizing individuals that traditionally underlie thinking about behavior (Loewenstein, Asch, Friedman, Melichar, & Volpp, 2012). Conventional economics considers (lifestyle) decisions as the outcome of an individual's information-based trade-off between costs and benefits (Boot & Van Lienden, 2011). Likewise, traditional psychological models (e.g. the Theory of Planned Behavior (Ajzen, 1985) or the Health Belief Model (Becker, 1974)) propose cognitively deliberated intentions as important triggers of behavior change.

In conventional models of behavior, self-control problems are typically denied; people do what they want and their actions reflect their preferences. This implies that self-damaging behaviors point to either a lack of knowledge or a lack of interest in a healthy lifestyle. This reasoning cannot be reconciled with introspection, contemporary psychological research and the observation of behaviors that point to self-control problems (Loewenstein et al., 2012).

An estimated 8.5 million Dutch people want to lose weight, while levels of obesity are rising (Kooreman & Prast, 2010) and year-long gym subscriptions are systematically overpaid and underused (DellaVigna & Malmendier, 2006). It appears that people know the benefits of change and genuinely intend to do so, but also procrastinate and have trouble acting on their intentions.

Contemporary behavioral science admits the limitations of human rationality

An ever expanding body of empirical evidence shows that people deviate from rationality in foreseeable situations. Without the assumptions of perfect rationality, behavioral science recognizes actual behavior and identifies systematic patterns in it (Prast, 2017). The acknowledgement of systemic deviations from the rational model as sources of behavior, such as emotions (Loewenstein & Lerner, 2003; Zeelenberg, Nelissen, Breugelmans, & Pieters, 2008), lapses of self-control (Thaler, 1981), and heuristics and biases (Tversky & Kahneman, 1974), points to a large class of circumstances in which people can be helped in achieving their own long-term goals (Loewenstein et al., 2012).

Self-control

One of the key contributions of psychology to economics (i.e., behavioral economics) is the insight that people have limited self-control (Ainslie, 1975), which harms their personal goals. Behavioral economic models of self-control accentuate time-inconsistency in judgment and decision-making. People balance costs and benefits differently over different time horizons (Soman et al., 2005): we tend to choose more deliberately when contemplating the future and more impulsively when choosing for the present. This dynamic inconsistency (Kirby & Herrnstein, 1995) has been dubbed present bias or metaphorically as the friction between the cold and farsighted *planner* and the 'hot' and myopic *doer* within us (Loewenstein, 2005; Thaler & Shefrin, 1981). Scientific models that incorporate self-control problems typically explain empirical observations of human behavior better than models that assume perfect rationality (Green & Myerson, 2010; Laibson, 1997).

Behavioral models of self-control project how people systematically overweigh the present. Not surprisingly, we generally desire good things sooner rather than later. Hence, waiting decreases the desirability of personal benefits: delayed outcomes are said to be discounted. In contrast to what is perfectly rational (Samuelson, 1937), the degree of delay discounting is not stable over time, but decreases as the length of the delay increases; a pattern known as hyperbolic discounting (Mazur, 1987). Hyperbolic discounting is a formalization of the fact that people generally dislike waiting *now* more than they expect to dislike waiting in the future (Laibson, 1997). We are not only impatient; we also underestimate how impatient we will be in the future and we overestimate our future self-control.
Looking forward, we may desire a healthy physique and enroll for a gym membership under the impression that the long-term benefits will outweigh the effort in the future, but later, in the

face of immediate gratification, we are more likely to watch more Netflix instead of exercising. Hence, hyperbolic time discounting can help explain self-control problems and accompanied procrastination of lifestyle improvement.

People are aware of difficulties and try to overcome them

From an intrapersonal perspective, present bias results in suboptimal health outcomes. Fortunately, similar to the awareness of the relation between health and behavior, people are not always unaware of their self-control difficulties. Within the context of foreseeing self-control problems, Laibson (1997) and O'Donoghue and Rabin (1999) have distinguished two extremes; sophisticates and naïfs. Sophisticates are described as individuals who foresee their self-control troubles and who may take measures to protect their long-term goals from their short-term (emotional) impulses. Naïfs also have self-control issues, but do not foresee their tendencies and may be unjustly convinced that they will stick to their goals. O'Donoguhue & Rabin (1999, p2.) describe the distinction as follows: "Intuitively, a sophisticated person is correctly pessimistic about her future behavior – a naïve person believes she will behave herself in the future while a sophisticated person knows she may not."

This PhD thesis focusses on the latter group; people who are aware of the benefits of a healthy lifestyle, who want to realize their personal health goals, but also feel that they may not act on this in the future. In different personal domains, people embrace or self-impose measures to circumvent future temptation. For example, some people literally freeze their credit cards in blocks of ice (Ariely, 2009) or impose withdrawal penalties on their savings accounts to avoid overspending (Beshears et al., 2015). Likewise, Dutch employees are happy with mandatory pension savings because "Otherwise I would not save enough" (Van Rooij, Kool, & Prast, 2007). Interviews with Dutch citizens also point to self-imposed choice restrictions as the most preferred strategy to maintain a healthy diet: "Do not take the tempting food into the house to begin with" (Van der Lucht & Polder, 2010).

These are all examples of people who recognize their limited self-control and feel that limiting their own future freedom might be beneficial in the long-run. Voluntary accepted restrictions on future decision-options are known as *commitment devices.* A common form of an effective commitment device is a deposit contract, via which people deposit their money and only get it back if they have lost weight (John et al., 2011), attended the gym (Goldhaber-Fiebert, Blumenkranz, & Garber, 2010), or abstained from smoking (Giné, Karlan, & Zinman, 2010) by a prespecified deadline. The costly deadlines are meant to circumvent self-control problems by drawing the consequences of procrastination nearer. As such, voluntary accepted deadlines with consequences strategically restrict future decision-options to facilitate goal-attainment.

Commitment lotteries

The focus of this PhD thesis is on commitment lotteries. Similar to deposit contracts, the lotteries are meant to assist people in preventing their self-control troubles. In a commitment lottery, participants set a behavioral health goal, to be achieved at a prespecified deadline. On the deadline, a prize is drawn out of all participants and announced to all. Importantly, the winners are only eligible for their prize if they attained their personal goal. As a consequence, non-eligible winners are informed about their forgone earnings. This counterfactual feedback is designed to provoke anticipated regret and emphasize the lottery deadlines.

In previous applications, similar lotteries in the United States of America have successfully supported medication adherence (Kimmel et al., 2012), weight loss (Volpp et al., 2008), and walking (Patel et al., 2016). Although their appeal is apparent, it remained unclear if- and in what context- this concept would be effective in the Netherlands. Besides, in previous instances, intervention effects were typically not maintained after an initial intervention period. The commitment lotteries that are discussed in this PhD thesis were innovated on multiple aspects in order to gain a better understanding of these open issues. For this thesis, it was studied whether commitment lotteries would support regular gym attendance in Dutch company gyms for up to 52 weeks and what (psychological) design features could contribute to long-term behavior change.

Objectives

The primary research questions of this PhD thesis were: what is the 1) short-term and 2) long-term effectiveness of different commitment lotteries in supporting lifestyle decisions that are in line with people's own goals? And 3) what are the contextual and psychological factors that help explain and optimize their effect and design?

Overview

In a commitment lottery, winners who have not attained their goals do not get their prize but receive feedback on what their forgone earnings would have been. This counterfactual feedback is designed to provoke anticipated regret and increase commitment to health goals. In Chapter 2, we explored in an experimental scenario-study which emotions were expected upon missing out on a prize and which incentive-characteristics influence their likelihood and intensity. These insights were used for Chapter 3, which describes the design and protocol of the 52-week cluster randomized trial to study if commitment lotteries would promote physical activity among overweight adults. Chapter 4 describes the execution and results of the short-term (13 weeks) and long-term (26 weeks) lottery-interventions until 26 weeks. In Chapter 5, the results of a 52-week follow-up and weight patterns over the course of the trial are described. In Chapter 6, the goal was to gain more insights into the design features that can help explain the effectiveness of commitment lotteries. In three experiments, it was explored to what extent counterfactual feedback retains its influence on decision-making if it is delayed and whether

feedback influences delay discounting. In the Discussion chapter, findings are reflected in light of scientific developments. Furthermore, the policy and practical implications of this PhD thesis are discussed.

REFERENCES

Ariely, D. (2009). *Predictably irrational*. New York: HarperCollins.

Acland, D., & Levy, M. R. (2015). Naiveté, projection bias, and habit formation in gym

Attendance. *Management Science, 61*(1), 146-160.

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin, 82*(4), 463- 496.

Ajzen I. (1985) From intentions to actions: a theory of planned behavior. In: Kuhl J., Beckmann J. (Eds.), *Action Control* (pp. 11-39). SSSP Springer series in social psychology. Springer: Berlin, Heidelberg.

Aminov, R. I. (2010). A brief history of the antibiotic era: lessons learned and challenges for the future. *Frontiers in Microbiology, 1* (134), 1-7.

Becker, M. H. (1974). The health belief model and sick role behavior. *Health Education Monographs, 2*(4), 409-419.

Beshears, J., Choi, J. J., Harris, C., Laibson, D., Madrian, B. C., & Sakong, J. (2015). Self control and commitment: can decreasing the liquidity of a savings account increase deposits? *National Bureau of Economic Research, w21474*, 1–53.

Bickel, W. K., Moody, L., & Higgins, S. T. (2016). Some current dimensions of the behavioral economics of health-related behavior change. *Preventive Medicine, 92* (C), 16-23.

Boot, J. M. D., & Van Lienden, H. W. (2011). *Economie van de volksgezondheid. zorg voor gezondheid = zorg voor welvaart*. Assen: Koninklijke Van Gorcum.

Carrera, M., Royer, H., Stehr, M., & Sydnor, J. (2018). Can financial incentives help people trying to establish new habits? Experimental evidence with new gym members. *Journal of Health Economics, 58*, 202-214.

CBS, Central Bureau voor de Statistiek. (2018). *Gezondheid, leefstijl, zorggebruik en -aanbod, doodsoorzaken; kerncijfers*. Retrieved from statline.cbs.nl.

CBS, Central Bureau voor de Statistiek., & RIVM, Rijksinstituut voor Volksgezondheid en Milieu. (2017). *Gezondheidsenquête/Leefstijlmonitor*. Retrieved from statline.cbs.nl.

CPNB ,Stichting Collectieve Propaganda van het Nederlandse Boek. (2018). *CPNB Top 100 2017*. Retrieved from https://www.cpnb.nl.

DellaVigna, S., & Malmendier, U. (2006). Paying not to go to the gym. *The American Economic Review, 96*, 694-719.

Elfhag, K., & Rössner, S. (2005). Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. *Obesity Reviews, 6*(1), 67-85.

European Commission: Strategy, Corporate Communication Actions and Eurobarometer Unit

(2014). Sport and physical activity. *Special Eurobarometer 412*.

Forouzanfar, M. H., Afshin, A., Alexander, L. T., Anderson, H. R., Bhutta, Z. A., Biryukov, S., . . . Charlson, F. J. (2016). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388*(10053), 1659-1724.

Giles, J. (2011). Social science lines up its biggest challenges. *Nature, 470*, 18-19.

Giné, X., Karlan, D., & Zinman, J. (2010). Put your money where your butt is: a commitment contract for smoking cessation. *American Economic Journal: Applied Economics*, *2* (4), 213-235.

Goldhaber-Fiebert, J. D., Blumenkranz, E., & Garber, A. M. (2010). Committing to exercise: contract design for virtuous habit formation. *National Bureau of Economic Research ,w 16624*,1-25.

Green, L., & Myerson, J. (2010). Experimental and correlational analyses of delay and probability discounting. In G. J. Madden & W. K. Bickel (Eds.), *Impulsivity: The behavioral and neurological science of discounting* (pp. 67-92). Washington, DC, US: American Psychological Association.

Hulsegge, G., van der Schouw, Y. T., Daviglus, M. L., Smit, H. A., & Verschuren, W. M. (2016). Determinants of attaining and maintaining a low cardiovascular risk profile-the Doetinchem Cohort Study. *European Journal of Public Health, 26*(1), 135-140.

ING Bank. (2017). *Goede voornemens 2018*. Retrieved from www.ing.nl/media/A3%20Goede%20 voornemens_tcm162-137795.pdf.

John, L., Loewenstein, G., Troxel, A., Norton, L., Fassbender, J., & Volpp, K. (2011). Financial incentives for extended weight loss: a randomized, controlled trial. *Journal of General Internal Medicine, 26*(6), 621-626.

Kimmel, S. E., Troxel, A. B., Loewenstein, G., Brensinger, C. M., Jaskowiak, J., Doshi, J. A., . . . Volpp, K. (2012). Randomized trial of lottery-based incentives to improve warfarin adherence. *American Heart Journal, 164*(2), 268-274.

Kirby, K. N., & Herrnstein, R. J. (1995). Preference reversals due to myopic discounting of delayed reward. *Psychological Science, 6*(2), 83-89.

Kooreman, P., & Prast, H. (2010). What does behavioral economics mean for policy? Challenges to savings and health policies in the netherlands. *De Economist, 158*(2).

KVB-SMB/GfK. (2017). *Aandelen op basis van GfK-panel van de Nederlandse A-boekenmarkt.* Ad hoc analyses obtained from GfK upon request.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112(2), 443-477.

Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet, 380* (9838), 219-229.

Loewenstein, G. (2005). Hot-cold empathy gaps and medical decision making. *Health Psychology, 24*(4S), S49-56.

Loewenstein, G., Asch, D. A., Friedman, J. Y., Melichar, L. A., & Volpp, K. G. (2012). Can behavioural economics make us healthier? *BMJ: British Medical Journal (Online), 344, e3482*.

Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. In R. Davidson, H. Goldsmith, & K. Scherer (Eds.), *Handbook of Affective Science* (pp. 619-642). Oxford: Oxford University Press.

Van den Berg, M. & Schoemaker, C.G. (2010). *Effecten van preventie. Deelrapport van de VTV 2010: Van gezond naar beter.* Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu.

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In Commons, M.L., Mazur, J.E., Nevin, J.A. (Eds), *Quantitative analyses of behavior- the effect of delay and of intervening events on reinforcement* (pp. 55-73). New Jersey: Lawrence Erlbaum Associates.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review, 89*(1), 103-124.

OECD and European Observatory on Health Systems and Policies (2017). *State of Health in the EU. Netherlands: Country Health Profile 2017*. Paris: OECD.

OECD Data (2015). *Life expectancy at birth*. Retrieved from: https://data.oecd.org/healthstat/life-expectancy-at-birth.htm.

WHO, World Health Organization. (2015). *Global Health Observatory data*

repository. Retrieved from http://apps.who.int/gho/data.

Patel, M. S., Asch, D. A., Rosin, R., Small, D. S., Bellamy, S. L., Eberbach, K., . . . Wesby, L. (2016). Individual versus team-based financial incentives to increase physical activity: a randomized, controlled trial. *Journal of General Internal Medicine, 31*, 1-9.

Prast, H. (2017). De psychologie van pensioenkeuzes. *Netspar, Network for Studies on Pensions, Aging and Retirement, Brief 10*.

RIVM, Rijksinstituut voor Volksgezondheid en Milieu. (2018). *Volksgezondheid Toekomst Verkenning 2018: Een gezond vooruitzicht - Synthese Leefstijl en omgeving*. Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu.

Rijksoverheid. (2018). Meer ruimte voor preventie in basispakket zorgverzekering [Press release]. Retrieved from www.rijksoverheid.nl/actueel/nieuws/2018/06/01/meer-ruimte-voor-preventie-in-basispakket-zorgverzekering.

Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies, 4*(2), 155-161.

Soman, D., Ainslie, G., Frederick, S., Li, X., Lynch, J., Moreau, P., . . . Trope, Y. (2005). The psychology of intertemporal discounting: Why are distant events valued differently from proximal ones? *Marketing Letters, 16*(3-4), 347-360.

Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters, 8*(3), 201-207.

Thaler, R. H., & Shefrin, H. M. (1981). An economic theory of self-control. *The Journal of Political Economy, 89* (2), 392-406.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

United Nations. (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*. New York: Sustainable development knowledge platform.

United Nations: General Assembly. (2010). *Resolution 64/265: Prevention and control of non-communicable diseases*. New York: 86th plenary meeting.

Van der Lucht, F., & Polder, J. J. (2010). *Van gezond naar beter. Volksgezondheid Toekomst Verkenning 2010*. Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu.

Van Rooij, M. C., Kool, C. J., & Prast, H. M. (2007). Risk-return preferences in the pension domain: are people able to choose? *Journal of Public Economics, 91*(3), 701-722.

Van Winkelhof, M., Pijl, H., & Vliegenthart, M. (2018). Gezondere levensstijl is een onderschat medicijn voor de moderne mens. *NRC*. Retrieved from www.nrc.nl/nieuws/2018/04/03/aanpak-van-leefstijlziekten-levert-meer-op-a1598039.

Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial incentive–based approaches for weight loss: A randomized trial. *JAMA, 300*(22), 2631-2637.

Van Zon, H. (1990). *Tachtig jaar RIVM*. Assen: Van Gorcum.

Zeelenberg, M., Nelissen, R. M., Breugelmans, S. M., & Pieters, R. (2008). On emotion specificity in decision making: Why feeling is for doing. *Judgment and Decision Making, 3*(1), 18-27.

CHAPTER 2

Emotional responses to behavioral economic incentives for health behavior change

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Zeelenberg, M., Polder, J. J., & Prast, H. M. (2018). Emotional responses to behavioral economic incentives for health behavior change. *Psychology, Health & Medicine 23*(8), 996-1005.

ABSTRACT

Many people aim to change their lifestyle, but have trouble acting on their intentions. Behavioral economic incentives and related emotions can support commitment to personal health goals, but the related emotions remain unexplored. In a *regret lottery*, winners who do not attain their health goals do not get their prize but receive feedback on what their forgone earnings would have been. This counterfactual feedback should provoke anticipated regret and increase commitment to health goals. We explored which emotions were actually expected upon missing out on a prize due to unsuccessful weight loss and which incentive-characteristics influence their *likelihood* and *intensity*. Participants reported their expected emotional response after missing out on a prize in one of 12 randomly presented incentive-scenarios, which varied in incentive type, incentive size and deadline distance. Participants primarily reported feeling disappointment, followed by regret. Regret was expected most when losing a lottery prize (vs. a fixed incentive) and intensified with prize size. Multiple features of the participant and the lottery incentive increased the occurrence and intensity of regret. As such, our findings can be helpful in designing behavioral economic incentives that leverage emotions to support health behavior change.

'The ultimate currency that rewards or punishes is often emotional'
-Daniel Kahneman (2011, p.343)

Currently, 62% of Europeans and 74% of Americans are overweight or obese (Flegal, Carroll, Kit, & Ogden, 2012; WHO, 2015). Consequently, one of the key challenges of the modern-day health professional is effectively supporting people who wish to improve their lifestyle.
A promising direction is the use of financial incentives for health behavior change (Mantzari et al., 2015). To improve their impact, behavioral economists have tested lotteries that are designed to leverage regret aversion (Volpp et al., 2008). Generally, people anticipate future regret if they expect to learn the outcome of a non-chosen opportunity (Zeelenberg, 1999). As such, regret can improve health decisions such as vaccination (Chapman and Coups, 2006; Lagoe and Farrar, 2015), use of contraceptives (Richard, De Vries, & Van der Pligt, 1998; Smerecnik and Ruiter, 2010), and exercising (Abraham and Sheeran, 2003).

Volpp et al., (2008) used the psychology of regret to optimize lottery-incentives that were designed to help people attain their weight loss goal. If participants won the lottery, they could only claim their prize if they had attained their predetermined weight loss goal. The winning ticket was drawn out of all participants and non-eligible lottery winners learned what their forgone earnings would have been (i.e. counterfactual feedback). A meta-analysis by Haff et al. (2015), evaluating multiple applications of the lotteries, targeted at various health behaviors, projected a pooled goal-attainment of 57.5%, opposed to 22.6% without lotteries (Haff et al., 2015).

Due to the counterfactual feedback in the lotteries, Haff and colleagues labeled the lotteries as *regret lotteries*. Likewise, in explaining the effectiveness of the lotteries, Volpp and colleagues stated that "the anticipated threat of regret" (p. 2636) could help explain why participants attained their weight loss goals. However, it remains unexplored which emotions are expected when missing out on a prize and which incentive-characteristics influence the *likelihood* and *intensity* of these emotions.
This is important knowledge because different emotions prompt different behaviors (Frijda, 1987, 2007) and logically, incentives that leverage emotions should commit goal-striving participants to goal directed behaviors (e.g. exercising). Besides, expected emotion intensity generally increases the likelihood of goal directed behavior (Frijda, 2007; Loewenstein and Lerner, 2003). Hence, exploring which incentive-characteristics contribute to which emotional responses can contribute to the further optimization of health incentives.

The current exploration has three aims. First, we explore which emotions are expected upon missing out on of 12 different incentives. Second, we explore which incentive-characteristics influence the *likelihood* of the reported emotions. Third, we explore the incentive-characteristics that contribute to the *intensity* of reported emotions.

METHOD

We described missing out on a prize in a hypothetical scenario of unsuccessful weight loss and asked participants to report their expected emotions and emotion intensity. We varied three basic incentive-characteristics that one needs to consider when designing an incentive to promote health behavior change (Adams, Giles, McColl, & Sniehotta, 2014; Halpern, Asch, & Volpp, 2012). The incentive-characteristics that were varied were incentive type, incentive size and deadline distance. As such, a 2 (lottery vs. fixed prize) × 3 (€50 vs. €500 vs. family vacation as prize) × 2 (6-month deadline vs. 12-month deadline) between-subjects scenario-design was used.

Participants

Data was collected through an internet survey among participants of the CentERpanel in the Netherlands. The CentERpanel consists of about 2000 households representative of the Dutch-speaking population in the Netherlands. Upon deciding to enter the CentERpanel, members are explained that their survey-responses will be used exclusively for non-commercial purposes. A total of 1369 participants between the ages of 18-65 were presented with a questionnaire. Fourteen participants were excluded due to not answering the questions and 26 were excluded because their commentary strongly indicated that they were not seriously participating. As such, the initial sample consisted of 1329 participants with a mean age of 46.4 (*SD* = 12.13) half of whom (51.9%) was female.

Procedure and Materials

All participants were asked to respond to one of 12 randomly presented scenarios in a questionnaire. All scenarios started as follows: "*Imagine that you have the goal to lose weight and that you are offered some assistance. Together with your health center you determine a 10-week target weight.*" The scenarios next systematically varied between-subjects in incentive type, size and deadline distance.

In the lottery scenarios, participants read the following text: *For commitment purposes, you are offered to participate in a free lottery with a prize of* (€50 or €500 or a family vacation)*. You can always win the lottery, but you can only claim your prize if you achieve your target weight after 10 weeks and remain at or below this weight at the* (6 or 12)*-month deadline. The winning ticket is drawn out of all participants and you always get feedback on the outcome of the lottery.* Participants were next asked to what degree they would be willing to participate (1 = not at all; 6 = very much).

In the fixed prize scenarios, participants read the following scenario: *For commitment purposes you are offered a reward of* (€50 or €500 or a family vacation) *if you achieve your target weight after 10 weeks and remain at or below this weight at the* (6 or 12)*-month deadline.* Participants were next asked to what degree they would be willing to participate (1 = not at all; 6 = very much).

Next, in the lottery scenarios, participants read the following text: *Now imagine that you win the lottery. Unfortunately you cannot claim your prize because you did not achieve your target weight.*
In the fixed prize scenarios, participants next read the following text: *Now imagine that you are at the deadline and you are not rewarded because you did not achieve your target weight.*

After reading one of 12 scenarios, participants were asked to select, out of 15 randomly presented emotions, the primary emotion that they would feel at this point. We next sequentially asked participants to select the second and third emotion they would feel (based on Zeelenberg and Pieters, 2004, who assessed lottery-based emotions; see Table 2). Participants were also asked to indicate to what degree they would feel the selected emotions (i.e. emotion intensity) and, if not selected, the degree of regret (1 = not at all; 6 = very intense). Participants were next asked to state their 'subjective' need to lose weight (weight loss intention; 0 = no; 1 = yes) and their current weight and height as an assessment for their 'objective' need to lose weight (BMI).

Finally, the five-item Regret Scale (RS, $\alpha = .84$; Schwartz et al., 2002) was presented to assess a personal tendency to compare decision-related outcomes. The validated RS is often used to measure regret proneness (e.g., Saffrey, Summerville, & Roese, 2008; Spunt, Rassin, & Epstein, 2009) and had the benefit of being short while being reliable and informative.

RESULTS

Descriptives

The mean score on willingness to participate was 3.67 (*SD* = 1.68). To increase the chance of the participants being able to truly imagine themselves in the presented scenario, subsequent analyses were performed among the subsample of participants who were willing to participate in a weight loss initiative. The central score (3 on a scale of 1 to 6) was used as a demarcation of high and low willingness. The high-willingness sample contained 763 participants (57.4%) with a willingness-score > 3 (see Table 1 for an overview), about half of whom was female (51.9%). The mean age was 45.05 (*SD* = 12.22) and mean BMI was 25.45 (*SD* = 4.15).

Table 1. Random allocation of 763 participants to one of 12 scenarios.

Lottery	**€ 50**	**€ 500**	**Vacation**
6 month-deadline	*n* = 61	*n* = 49	*n* = 51
12 month-deadline	*n* = 43	*n*= 54	*n* = 57
Fixed incentive	**€ 50**	**€ 500**	**Vacation**
6 month-deadline	*n* = 62	*n* = 88	*n* = 76
12 month-deadline	*n* = 58	*n* = 78	*n* = 86

Example: scenario 1 described a lottery with a €50 prize and a 6-month deadline.

Table 2 provides an overview of the stated emotions. Six emotions were mentioned by more than 20% of participants and were considered for further analysis. In the lottery scenarios, 76.5% expected feeling disappointment and 51.7% of the participants expected feeling regret when missing out on their prize. A total of 24.9% stated feeling both regret and disappointment (first, second or third mentioned emotion) when deprived of their prize. Guilt was reported by 40.3% and 28% expected feeling shame. Irritation was expected by 29.5% of participants and sadness by 22.5%.

In the scenarios that described being withheld of a fixed incentive, 82.2% expected feeling disappointment and 46.5% expected feeling regret. Guilt was reported by 41.1% of participants and 35.5% expected irritation. Shame and sadness were reported by 27% and 24.5% respectively when missing out on a fixed incentive.

Table 2. Stated emotions when missing out on a prize

	First Emotion		Second Emotion		Third Emotion		Total	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
Lottery								
Disappointment	157	49.8	58	18.4	26	8.3	241	76.5
Regret	52	16.5	64	20.3	47	14.9	163	51.7
Guilt	27	8.6	46	14.6	54	17.1	127	40.3
Shame	26	8.3	26	8.3	36	11.4	88	28
Sadness	16	5.1	30	9.5	25	7.9	71	22.5
Irritation	11	3.5	40	12.7	42	13.3	93	29.5
Anger	7	2.2	15	4.8	17	5.4	39	12.4
Pride	7	2.2	10	3.2	8	2.5	25	7.9
Relief	5	1.6	8	2.5	18	5.7	31	9.8
Happiness	3	1	6	1.9	8	2.5	17	5.4
Jealousy	3	1	2	0.6	4	1.3	9	2.9
Disgust	1	0.3	3	1	9	2.9	13	4.2
Envy	0	0	4	1.3	11	3.5	15	4.8
Fear	0	0	1	0.3	0	0	1	0.3
Elation	0	0	2	0.6	10	3.2	12	3.8
Fixed								
Disappointment	279	62.3	60	13.4	29	6.5	368	82.2
Guilt	42	9.4	76	17	66	14.7	184	41.1
Regret	35	7.8	96	21.4	79	17.6	210	46.8
Shame	22	4.9	45	10	54	12.1	121	27
Irritation	20	4.5	65	14.5	74	16.5	159	35.5
Sadness	19	4.2	39	8.7	52	11.6	110	24.5
Pride	8	1.8	5	1.1	14	3.1	27	6
Anger	7	1.6	26	5.8	23	5.1	56	12.5
Relief	7	1.6	7	1.6	27	6	41	9.2
Disgust	4	0.9	6	1.3	6	1.3	16	3.5
Elation	3	0.7	11	2.5	5	1.1	19	4.3
Fear	1	0.2	2	0.4	2	0.4	5	1
Happiness	1	0.2	8	1.8	6	1.3	15	3.3
Envy	0	0	0	0	7	1.6	7	1.6
Jealousy	0	0	2	0.4	4	0.9	6	1.3

Note: Participants were sequentially asked to state their first, second and third emotional response to a lost prize.

Likelihood of emotions

Logistic regression analyses were performed to determine the incentive-characteristics that contribute to the likelihood of the emotions. Only the first-chosen emotion was used (0 = not mentioned first, 1 = mentioned first), so that the model would distinctively predict the emotion of interest.

In six independent analyses, disappointment, regret, guilt, shame, sadness and irritation were used as dependent variables respectively. The incentive-characteristics from the scenarios, age, sex, the RS, objective- and subjective need to lose weight were entered as independent variables. None of the incentive-characteristics significantly influenced the likelihood of guilt, shame, sadness or irritation (results not further displayed). In contrast, the likelihood of regret and disappointment was influenced by the incentive-characteristics and therefore reported in Table 3.

Table 3. Characteristics influencing the likelihood of Regret and Disappointment, logistic regression.

	Regret			Disappointment		
	OR	95% C.I.		OR	95% C.I.	
		Lower	Upper		Lower	Upper
Lottery vs. Fixed	2.55**	1.44	4.52	0.63*	0.44	0.92
12 vs. 6 months	1.04	0.59	1.82	0.94	0.65	1.36
€500 vs. €50	2.10	0.97	4.54	0.82	0.52	1.29
Vacation vs. €50	2.42*	1.13	5.21	0.87	0.55	1.37
Age	1.08	0.80	1.45	1.00	0.82	1.21
Female vs. Male	1.02	0.58	1.79	0.89	0.62	1.29
BMI	0.96	0.69	1.35	0.89	0.72	1.11
Intention	2.07*	1.09	3.95	0.95	0.63	1.44
Regret Scale	0.87	0.65	1.17	0.88	0.72	1.06
Constant	0.04			1.85		

Nagelkerke R^2: Regret =.09. Disappointment =.03
Cox & Snell R^2: Regret = .05. Disappointment = .02
*Significant at $p < .05$
**Significant at $p < .01$

Missing out on the lottery prize elicited regret significantly more often than being deprived of the fixed incentive (OR = 2.55, $p = .001$, 95% CI, 1.44 to 4.52). Losing the vacation (vs. €50) also significantly increased the likelihood of regret (OR = 2.42, $p = .02$, 95% CI, 1.13 to 5.21), and losing €500 (vs. €50) did not significantly increase the likelihood of regret at $p < 0.05$ (OR = 2.10, $p = .06$, 95% CI 0.97 to 4.54).
The objective need to lose weight (BMI) did not yield a significant parameter in predicting regret (OR = 0.96, $p = .82$, 95% CI, 0.69 to 1.35), whereas the subjective need to lose weight (intention) lead to a higher frequency of reported regret (OR = 2.07, $p = .03$, 95% CI, 1.09 to 3.95). The likelihood of disappointment increased when the incentive was fixed opposed to a lottery (OR = 0.63, $p = .02$, 95% CI, 0.44 to 0.92).

Intensity of emotions

Six independent linear regression analyses were performed to assess the different incentive-characteristics that intensify the emotions. The intensity of the emotions was used as dependent variable. The incentive-characteristics, age, sex, objective- and subjective need to lose weight and the RS were entered as independent variables.

None of the incentive-characteristics significantly influenced the intensity of guilt, shame or irritation. The intensity of sadness increased significantly as a result of losing a family vacation (B = .41, SE = .20, p = .04).

Results of the regressions of disappointment and regret are displayed in Table 4. The intensity of regret increased significantly when the lost incentive was lottery-based opposed to fixed (B = .35, SE = .13, p < .01). Regret also intensified when the prize was €500 (B = .37, SE = .16, p = .02) or a vacation, (B = .48, SE = .16, p < .01). Women (B = .43, SE = .13, p < .01) and participants with a personal proneness to feel regret (B = .16, SE = .07, p = .01) further expected feeling more intense regret.

Table 4. Characteristics influencing the intensity of Regret and Disappointment, OLS regression.

	Regret		Disappointment	
	B	S.E.	B	S.E.
Lottery vs. Fixed	0.35**	0.13	-0.21*	0.10
12 vs. 6 months	-0.07	0.12	-0.12	0.10
€500 vs. €50	0.37*	0.16	0.10	0.12
Vacation vs. €50	0.48**	0.16	0.12	0.12
Age	0.04	0.07	-0.18**	0.05
Female vs. Male	0.43**	0.13	0.30**	0.10
BMI	-0.04	0.02	0.00	0.06
Intention	0.19	0.14	0.30**	0.11
Regret Scale	0.16*	0.07	0.03	0.05
Constant	3.20	0.15	4.83	0.12

R^2: Regret = .08 Disappointment = .10
*Significant at p < .05
**Significant at p < .01

The intensity of disappointment increased when the incentive was fixed opposed to lottery-based (B = -0.21, SE = .10, p = .04). Incentive size did not significantly affect the intensity of disappointment. Additionally, women (B = .30, SE = .10, p < .01), relatively younger participants (B = -0.18, SE = .05, p < .01) and participants who intended to lose weight (B = .30, SE = .11, p < .01) reported more intense disappointment when missing out on their prize.

DISCUSSION

The aims of the current study were to explore 1) which emotions would be expected upon missing out on a prize and which incentive-characteristics would contribute to the 2) likelihood and 3) intensity of reported emotions. After reading one of 12 incentive-scenarios, participants primarily report feelings of disappointment and regret when missing out on a prize and to a lesser extent irritation, guilt, shame and sadness.

Regret

A lottery design (versus a traditional fixed incentive) increased the likelihood and intensity of regret, which helps substantiate the label *regret lotteries*. Besides, the expected intensity of regret increased with both increases in size of the loss and the likelihood of regret increased if participants imagined losing a family vacation (vs. €50). This pattern is in line with economic regret theory (Bell, 1982), in which regret is described as the discrepancy between the current situation and 'what would have been'. As such, a higher discrepancy results in more regret. The finding that losing a family vacation increases the likelihood and intensity of regret can also be interpreted in line with regret literature by Janis and Mann (1977) and Zeelenberg (1999) who theorized that socially important outcomes can intensify regret along with a simple increase in size of a bad outcome.

Deadline distance did not affect the likelihood or intensity of regret. This mirrors results from a meta-analysis in which inaction-regret influences behavior independent of the distance of the negative outcome (Brewer, DeFrank, & Gilkey, 2016). Still, it remains an interesting open question if deadline distance does not matter for incentives to evoke expectations of future regret and decision-making in field settings.

Participants-characteristics were also found to influence expected regret. The subjective need to realize weight-loss appears more relevant in eliciting regret than an objective need to lose weight: people who intend to lose weight, experience regret sooner (and more intense disappointment), whereas people with a higher BMI do not. This finding resembles the function of emotions as personal indicators of goal importance (Frijda, 2007; Zeelenberg, Nelissen, Breugelmans, & Pieters, 2008) and as such supports the idea that emotions can be used for goal commitment.

Regret and Disappointment

Regret is a universal emotion, experienced similarly across different cultures and while it is related to disappointment, the emotions also have some distinct antecedents and consequences (Breugelmans, Zeelenberg, Gilovich, Huang, & Shani, 2014). Disappointment is an emotional reaction to disconfirmed expectations (Bell, 1985; Loomes and Sugden, 1986). As regret, disappointment is a counterfactual emotion and can arise when comparing one's current situation to 'what could have been'. However, in the evaluation of a negative outcome, regret is more closely related to self-agency than disappointment (Frijda, 1987). Thus, people who feel regret feel more responsible for their bad situation than disappointed people. Therefore, it has been argued that disappointment is a broader response to an unfavorable outcome than regret (Zeelenberg, Van Dijk, Manstead, & Van der Pligt, 1998). Put differently, regret is a more centered emotion and stems from the realization that a disappointing outcome resulted from one's own behavior.

Our findings reflect this reasoning by demonstrating that the situational conditions leading to more (intense) regret are more specific than those that result in disappointment. Disappointment is reported broadly, while regret increases in more specific incentive-conditions. The present findings may therefore help in designing incentives that aim to leverage regret aversion.

Another feature that could help explain why disappointment is reported broadly in the current study is the contingency of the prize. For someone who did not achieve a target weight, reflecting on decisions that contributed to this outcome may be difficult because weight loss is no single decision, but a delayed outcome of a sequence of decisions.

We mainly focused on the characteristics of the incentive and not the target outcome. Future research could extend our findings by also varying the target outcome (e.g., gym attendance versus food intake) and reveal whether a lottery prize contingent on a specific behavior influences emotional responses to a loss.

A limitation of this study is that participants were asked to report (the intensity of) their expected emotions, but did not have to make an actual decision. We aimed to increase the practical relevance of our findings by selecting the subsample of participants who would actually be willing to participate in the presented program and by controlling for multiple covariates.

Although ample research has shown that expected (intensity of) emotions influence(s) decision-making (Frijda, 2007; Loewenstein and Lerner, 2003; Zeelenberg and Pieters, 2004), it remains uncertain if participants in our study would also act on their expected negative emotions.

Conclusion

Emotions can improve the effectiveness of health incentives (Haff et al., 2015). Therefore, it can be useful to have an indication of the emotional responses to different incentive designs. We explored emotional responses to missing out on a prize due to unsatisfactory weight loss, previously presented as *regret lotteries*. Disappointment is broadly experienced and several aspects of the participant and the lottery incentive were found to increase the occurrence and intensity of regret. The present findings may be helpful in designing lottery-based commitment programs to promote health behavior change. More research on the behavioral contingency of the prize would further improve the potential for effective commitment.

REFERENCES

Abraham, C., & Sheeran, P. (2003). Acting on intentions: The role of anticipated regret. *British Journal of Social Psychology, 42*(4), 495-512.

Adams, J., Giles, E. L., McColl, E., & Sniehotta, F. F. (2014). Carrots, sticks and health behaviours: a framework for documenting the complexity of financial incentive interventions to change health behaviours. *Health Psychology Review, 8*(3), 286-295.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research, 30*(5), 961-981.

Bell, D. E. (1985). Disappointment in Decision Making Under Uncertainty. *Operations Research, 33*(1), 1-27.

Breugelmans, S. M., Zeelenberg, M., Gilovich, T., Huang, W.-H., & Shani, Y. (2014). Generality and cultural variation in the experience of regret. *Emotion*, 14(6), 1037- 1048.

Brewer, N. T., DeFrank, J. T., & Gilkey, M. B. (2016). Anticipated regret and health behavior: A meta-analysis. *Health Psychology, 35*(11), 1264-1275.

Chapman, G. B., & Coups, E. J. (2006). Emotions and preventive health behavior: worry, regret, and influenza vaccination. *Health Psychology, 25*(1), 82.

Flegal, K. M., Carroll, M. D., Kit, B. K., & Ogden, C. L. (2012). Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. *JAMA, 307*(5), 491-497.

Frijda, N. H. (1987). Emotion, cognitive structure, and action tendency. *Cognition and Emotion*, 1(2), 115-143.

Frijda, N. H. (2007). *The laws of emotion*. Mahwah: Lawrence Erlbaum Associates Publishers.

Haff, N., Patel, M. S., Lim, R., Zhu, J., Troxel, A. B., Asch, D. A., & Volpp, K. G. (2015). The role of behavioral economic incentive design and demographic characteristics in financial incentive-based approaches to changing health behaviors: a meta-analysis. *American Journal of Health Promotion, 29*(5), 314-323.

Halpern, S. D., Asch, D. A., & Volpp, K. G. (2012). Commitment contracts as a way to health. *BMJ: British Medical Journal, 344*(1), e522-e522. .

Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. New York: Collier Macmillan.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Penguin.

Lagoe, C., & Farrar, K. M. (2015). Are you willing to risk it? The relationship between risk, regret, and vaccination intent. *Psychology, Health & Medicine, 20*(1), 18-24.

Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. In R. Davidson, H. Goldsmith, & K. Scherer (Ed.), *Handbook of Affective Science* (pp. 619-642). Oxford: Oxford University Press.

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies, 53*(2), 271-282.

Mantzari, E., Vogt, F., Shemilt, I., Wei, Y., Higgins, J. P., & Marteau, T. M. (2015). Personal financial incentives for changing habitual health-related behaviors: A systematic review and meta-analysis. *Preventive Medicine, 75*, 75-85.

Richard, R., De Vries, N. K., & Van der Pligt, J. (1998). Anticipated regret and precautionary sexual behavior. *Journal of Applied Social Psychology, 28*(15), 1411-1428.

Saffrey, C., Summerville, A., & Roese, N. J. (2008). Praise for regret: People value regret above other negative emotions. *Motivation and Emotion, 32*(1), 46-54.

Smerecnik, C. M. R., & Ruiter, R. A. C. (2010). Fear appeals in HIV prevention: The role of anticipated regret. *Psychology, Health & Medicine, 15*(5), 550-559.

Spunt, R. P., Rassin, E., & Epstein, L. M. (2009). Aversive and avoidant indecisiveness: Roles for regret proneness, maximization, and BIS/BAS sensitivities. *Personality and Individual Differences, 47*(4), 256-261.

Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial incentive–based approaches for weight loss: A randomized trial. *JAMA, 300*(22), 2631-2637.

WHO, World Health Organization. (2015). *Global Health Observatory data*

repository. Retrieved from http://apps.who.int/gho/data.

Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making, 12*(2), 93-106.

Zeelenberg, M., Nelissen, R. M., Breugelmans, S. M., & Pieters, R. (2008). On emotion specificity in decision making: Why feeling is for doing. *Judgment and Decision making, 3*(1), 18-27.

Zeelenberg, M., & Pieters, R. (2004). Consequences of regret aversion in real life: The case of the Dutch postcode lottery. *Organizational Behavior and Human Decision Processes, 93*(2), 155-168.

Zeelenberg, M., Van Dijk, W.W., Manstead, A. S. R., & Van der Pligt, J. (1998). The experience of regret and disappointment. *Cognition & Emotion, 12*(2), 221-230.

CHAPTER 3

Design and protocol of commitment lotteries: a cluster randomized trial

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Schipper, M., Zeelenberg, M., Polder, J. J., & Prast, H. M. (2016). Design and protocol of the weight loss lottery- a cluster randomized trial. *Contemporary Clinical Trials*, 49, 109-115.

ABSTRACT

People often intend to exercise but find it difficult to attend their gyms on a regular basis. At times, people seek and accept deadlines with consequences to realize their own goals (i.e., commitment devices). The aim of our cluster randomized controlled trial is to test whether a lottery-based commitment device can promote regular gym attendance. The winners of the lottery always get feedback on the outcome but can only claim their prize if they attended their gyms on a regular basis. In this paper we present the design and baseline characteristics of a three-arm trial which is performed with 163 overweight participants in six in-company fitness centers in the Netherlands.

People often intend to engage in physical activity (PA) on a regular basis, but have trouble putting their intentions into long-term behavior (DellaVigna & Malmendier, 2006). At times, people foresee their self-control difficulties and voluntarily elect arrangements that impede undesired future choices and actions, known as commitment devices (Rogers et al., 2014) Common applications of commitment devices are voluntarily depositing money into an account that can only be withdrawn upon goal-attainment, or making gym-appointments with a friend, where the cost of nonattendance is breaking a promise (Bryan, Karlan, & Nelson, 2010; Rogers et al., 2014).

Commitment devices for PA are especially beneficial for overweight individuals for multiple reasons. First, although physical unfitness is a hazard to individual health in all BMI-ranges (Barry et al., 2014), overweight individuals generally exercise less often than normal-weight individuals (CBS & RIVM, 2017), while PA can contribute to weight management (Fogelholm & Kukkonen-Harjula, 2000) and increase cardiorespiratory fitness, hereby reducing risks of numerous diseases (Hulsegge et al., 2015; Lee et al., 2012). Second, overweight and obesity have been associated with a relatively high dispositional desire for immediate gratification (Schlam et al., 2013; Tsukayama et al., 2010). Third, obesity lowers well-being especially among individuals with low self-control (Stutzer & Meier, 2015), while commitment devices may be helpful in overcoming self-control issues.

To investigate how overweight individuals can be helped in attaining their exercise goals, we test a lottery-based commitment device based on Volpp et al. (2008). Lottery participants set a gym-attendance goal and are handed multiple costly deadlines. At each deadline, the winning lottery ticket is drawn out of all participants. The winning participants, however, can only claim their prize if they attained their goal. Importantly, winning but nonattending participants are informed about their forgone earnings and thus know that they would have had a prize, had they attained their goal.

By promising explicit feedback on 'what would have been' at each deadline, we expect that participants anticipate feeling regret when missing out on their prize (Zeelenberg & Pieters, 2004). It is expected that anticipated regret of missing out on a lottery prize will commit people to their exercising goals (also see: Frijda, 2007; Zeelenberg et al., 2008).

Aim

The aim of this three-arm trial is to empirically study whether deadlines with lotteries can help people attain their goal of exercising twice a week. Additionally, we study whether weekly short-term lotteries for 13 weeks yield a different result than a long-term lottery after 26 weeks. Our secondary objective is to determine whether the intervention will result in physical and psychological changes.

We hypothesize that participants in both lottery-arms will attain their goals in more weeks than participants in the control-arm. We also hypothesize that after 26 weeks and at 52- week follow-up, participants in the long-term lottery-arm have attained their goal in more weeks than participants in the other arms. We further expect to observe a decline in goal attainment after removal of the lottery deadlines, but to a level above baseline. At follow-up, we hypothesize that goal attainment will be highest in the long-term lottery-arm and higher than control in the short-term lottery-arm.

METHOD

Setting

The Netherlands

The Commitment Lotteries are conducted in the Netherlands. Although the vast majority (92%) of the population knows that regular physical activity is important for a healthy life (Hildebrandt et al., 2007) a significant part (43%) of the Dutch population exercises less than once per week and nearly half is overweight (CBS & RIVM, 2017). Furthermore, approximately half of the population wants to lose weight (Kooreman & Prast, 2010). Accordingly, people intend to exercise regularly in the future, but the majority fails to follow through (Boshove, 2014). The preceding context highlights the need for (policy) initiatives by which the Dutch can effectively commit to their intentions, which would also benefit public health[1].

High Five

For this trial we cooperate with the international corporate gym enterprise High Five. High Five offers in-company fitness in 36 organizations across the Netherlands. From a convenience sample, six randomly selected gyms were approached to partake in our trial. The gyms were eligible to participate when their managers expressed their interest prior to randomization and if the managers were willing to invest time in scientific research. All of the six approached gyms met our eligibility criteria. Table 1 provides an overview of the involved business sectors. All gyms are run by gym managers who supervise several fitness coaches (or instructors). The coaches supervise the members, arrange fitness instructions and facilitate group classes.

1 This is not to say that a gap between intentions and behavior is a typical Dutch phenomenon. The discrepancy has been found in, for example, broader Europe (Biddle & Fuchs, 2009), Canada (Sniehotta et al., 2005) and the USA (DellaVigna & Malmendier, 2006).

Table 1. Study flow and treatment differences over time.

Study arm	Gym	Business sector	weeks 1-13		week 26		week 52
			weekly weigh-ins	weekly lottery	grand lottery	follow-up	follow-up
Control	1	Technology research	▲			▲	▲
	2	Municipality	▲			▲	▲
Short-term lotteries	3	Plastic fabrication	▲	▲		▲	▲
	4	Public sector research	▲	▲		▲	▲
Long-term lottery	5	Insurances	▲	▲	▲	▲	▲
	6	Pension administration & Investment	▲	▲	▲	▲	▲

Eligibility

The commitment device is studied with overweight participants ($25 \leq BMI < 40$) between the ages of 18-65 who explicitly stated to have the goal to exercise twice a week. Upper-BMI and age restrictions were used because participants outside these ranges generally require more consultation and supervision by the coaches, which may influence results. Participants had to be(come) members of one of the six participating gyms. Candidates were not eligible if they had planned a leave of absence of more than 4 weeks in the first 26 weeks of the trial. Including participants who violate this rule would, even before the start of the trial, disqualify them for a prize in one of the intervention arms (see below).

Interventions

The trial consists of two intervention arms and one control-arm. The American College of Sports Medicine and the American Heart Association recommend vigorous exercise for 20 minutes, three days a week, and muscular strength and endurance training two days a week (Haskell et al., 2007). Because all participating gyms are closed in the weekends, setting the goal of attending the gym two days a week was considered beneficial, while challenging but attainable. Hence, participants in all arms set the goal to attend their gym twice a week (the week-goal) and were handed a randomly generated three-digit study-ID prior to the start of the trial. With this ID, participants are required to register their attendance on iPads, provided to the six gyms. The regular attendance monitoring by High Five serves as a back-up. All participants are offered a monthly overview of their attendance via email.

Importantly, participants in both intervention arms are fully informed and reminded about the course and rules of the lotteries. Therefore, participants know that their number (study-ID) is in every drawing and that the outcome of the lottery is always communicated to them, regardless of their success.

Intervention arm 1: short-term lotteries

Free of charge, participants in this arm participate in a weekly lottery for 13 weeks with a weekly prize of €100. The winning study-ID is drawn out of all participants in this arm and the winners only obtain their prize if they have attended the gym at least twice that week. Any winner is informed about his or her earnings via email and text message. Importantly, lottery winners who did not attain their week-goal are informed about their forgone earnings. All other participants get to know whether the week-prize is awarded or not, but not to whom. Noteworthy is the fact that every new week offers a new opportunity to win and to keep attaining exercise goals, regardless of prior success. This rule fits a human inclination to use temporal landmarks (e.g. January or even Mondays) as a fresh start by relegating misfortune to the past (Dai et al., 2014). The weekly expected value for a compliant subject (i.e., exercising twice) is 1/60 × €100 = €1.67.

Intervention arm 2: long-term lottery

Intervention arm 2 is the same as the above-mentioned arm for the first 13 weeks. The weekly expected value for a compliant subject in the first 13 weeks in this arm is 1/55 × €100 = €1.82. Behavioral economic commitment schemes generally result in behavior change in the short-run, but the long-term effects are frequently unsatisfactory (Royer, Stehr, & Sydnor, 2012; Patel et al., 2016; Volpp et al., 2008). Setting a long-term meaningful deadline may overcome this issue, but people generally overvalue immediate costs (e.g. exercising) and discount future benefits, which may decrease the short-term influence of a long-term deadline (Laibson, 1997; O'Donoghue & Rabin, 1999). However, potential regret in the future has the ability to change behavior in the present by emphasizing the future consequences of current decisions (Richard, De Vries, & Van der Pligt, 1998). If, as we hypothesize, the promise of explicit counterfactual feedback about a lottery outcome is able to evoke anticipated regret, a long-term lottery may contribute to regular exercise patterns in the present, by emphasizing the consequences of not meeting the long-term deadline.

Therefore, in this arm, week 14-26 will also be part of the intervention. After week 26, we raffle off a luxury dream vacation for the participant and four friends or family members. The winning number is drawn out of all participants and communicated to the winner via email and via text message. The prize will only be awarded if the lottery winner has attained the week goal in at least 9 of the second 13 weeks (70% in weeks 14-26). If the winner does not meet the requirements for obtaining the prize, he or she receives a small consolation prize and another number is drawn until the prize can be claimed.

Control-arm

Participants in the control-arm also set the goal to attend the gym twice a week and are also monitored. They are also offered monthly statistics on their weight and attendance, but no lottery is organized. This way, the only designed treatment differences between intervention and control-arms are the lotteries.

Outcomes

Our primary behavioral outcome is the frequency of successful weeks after 13, 26 and 52 weeks in the study. A successful week is defined as having attained the week-goal (attended the gym at least twice).

Secondary outcomes of interest are the absolute frequency of gym visits, weight, fat percentage, fitness, mental well-being, exercising motivation and perceived health status (see Table 2 for all measures). Mental well-being and perceived health status are assessed to explore whether regular PA will affect both (see: Hassmén et al., 2000; Kettunen et al., 2015). Exercising motivation is assessed to explore whether any treatment effect is associated with the type of exercising motivation (e.g. external or intrinsic).

Table 2. Measurements over time.

Measurement	Construct	Baseline	13 weeks	26 weeks	52 weeks
High Five attendance records Self-report	Historic gym attendance	▲	▲	▲	▲
iPad self-registration [a]	Gym attendance		▲	▲	▲
KERN™ Scale [b]	Weight (kg), fat percentage	▲	▲	▲	▲
Physical Activity Rating & Non-Exercise Fitness Test (PA-R & NEFT; Jackson, 1990)	Cardiorespatory fitness using estimated VO_2max	▲	▲	▲	▲
Behavioral Regulation In Exercise Questionnaire (BREQ 2; Markland & Tobin, 2004)	Behavioural regulation in exercise	▲	▲	▲	▲
Warwick-Edinburgh Mental Well-being Scale (WEMWBS; Tennant et al., 2007)	Mental well-being	▲	▲	▲	▲
Regret Scale (Schwartz et al., 2002)	Regret tendency	▲			
Regret Experience Measure (REM; Creyer & Ross, 1999)	Anticipated regret		▲	▲	
Nationality		▲			
Education		▲			
Income		▲			

[a] Gym attendance is measured throughout the trial
[b] Weight is also measured weekly

Sample size

In a meta-analysis, Haff et al. (2015) evaluated similar lottery studies targeted at various health behaviors (e.g. weight loss, medication adherence), which projected a pooled success percentage of 57.5% in the lottery-arms versus 22.6% in the control-arms. A sample size calculation for a 0.35 difference between proportions at $p<.05$ and a power (1-β) of 80%, indicated a sample size of 27 per arm. With an intra-class correlation coefficient of .012, based on cluster research assessing PA at work (Gulliford et al., 1999), a design effect of 1.31 and an effective sample size of 36 per arm was estimated. We aimed to recruit a minimum of 50 participants per arm, allowing 25-30% attrition after 52 weeks.

Randomization

The six participating gyms (or clusters) were informed about the trial and agreed to participate prior to randomization. No within-gym randomization was performed to maintain participant blinding, avoid treatment contamination and to maximize the administrative convenience for the gym personnel. Therefore, every trial-arm contains two participating gyms.

Based on anonymized member data, we were able to distinguish three gyms with a relatively high percentage of overweight members and three gyms with a relatively low percentage of overweight members. By computer generation, first high-percentage gyms and next low-percentage gyms were randomly allocated to either one of the intervention arms or the control-arm. By this sequential two-step randomization, we aimed to avoid large differences in enrolment tempo.

Blinding

Participants were blinded from the treatments in other trial-arms. The researchers and gym personnel could not be blinded due to the design and comprehensive multi-party coordination of the study. Moreover, scholars (Given et al., 1990) and gym managers have stated that commitment by the staff can increase if they are treated as being part of the research team rather than being withheld from important aspects of the study. The involved managers and coaches were instructed not to inform the participants about the different treatments.

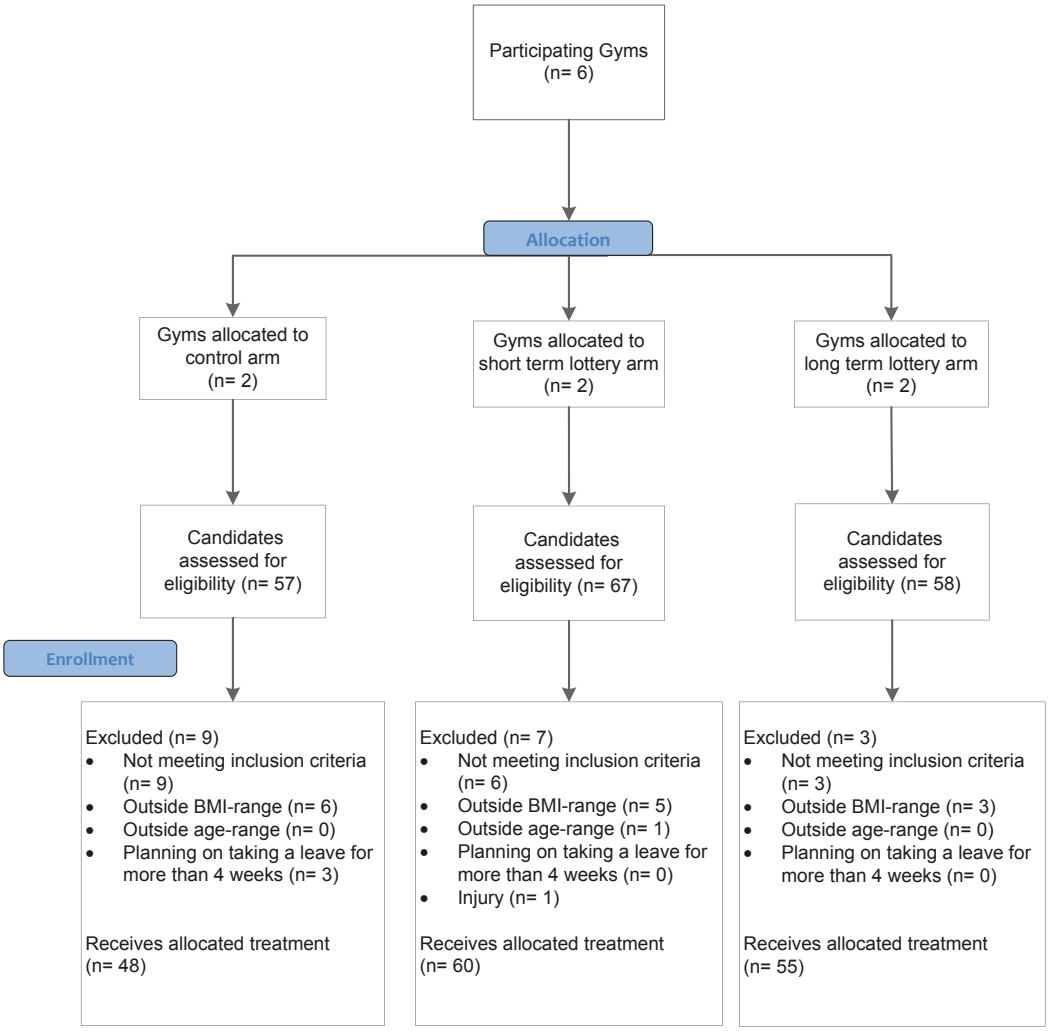
Instruction & recruitment

Next to written information, a study briefing was performed in all gyms. Additionally, we recorded an instruction video for all involved gym personnel. The video was offered to simplify future reference, enhance commitment to the study and remembrance of the study procedures. Existing gym members, as well as new members were eligible to participate in the trial. We aimed to recruit a minimum of 25 participants per gym, but allowed gyms to screen more participants. Gym managers were provided with a standard recruiting text, which explicitly targeted candidates who were looking for a commitment device for regular exercise. Managers communicated the text via email, in-company webpages, and newsletters. Additionally, posters, flyers and digital banners were designed and displayed by High Five.

Screening & enrollment

All communication materials contained a summary of the nature and procedure of the study and directed candidates to the fitness coaches. An online application assessing the eligibility criteria (see above) was developed and installed on iPads, provided to the gyms. Together, the candidate and coach could enter the candidate's characteristics. Based on the input, the application computed and checked the BMI, verified the other eligibility criteria and immediately specified whether the candidate was eligible for participation in the study. If so, informed consent was secured and participants were informed to wait for definitive admission. All application entries were manually verified by a member of the research team. The enrolment flow is displayed in Figure 1.

Figure 1. CONSORT Flow diagram of participant screening and enrollment.



Measures

The primary outcome is measured from an intent-to-treat perspective and is the frequency of successful weeks (measured binary: 1 = has attained week-goal 0 = has not attained week-goal). Gym attendance prior to the trial was assessed via self-report and will be extracted from the attendance register of every gym. During the trial, gym attendance is monitored by requiring participants to check-in via an online application on the iPad when entering their gym. Via the application, we are able to monitor attendance daily in our online database without the gym-personnel having to extract attendance data from their register every week.

For secondary outcomes, multiple measurements are performed. Weight and fat percentage are measured in the fitness centers. Paramount to our weighing protocol is that the measurements are as similarly as possible at all locations. Therefore, all gyms were provided with the same professional scales (KERN™; 0.1% precision). The assessments are highly similar to the existing corporate protocol; participants remove their shoes prior to weighing, after which fat percentage is determined by the scale with bioelectrical impedance analysis. The participants fill out the results on the iPad in their gym. At baseline and after 13, 26 and 52 weeks, participants are supervised by a gym-coach in entering these measures. Using online questionnaires, we assessed fitness, education level, income, nationality and psychological constructs.

Fitness is determined via the Physical Activity Rating & Non-Exercise Fitness Test (PA-R & NEFT; Jackson et al., 1990). PA-R & NEFT provide the opportunity to reliably estimate maximum oxygen uptake (VO_2max) with a questionnaire (0 "I avoid walking or exercise" – 7 "I regularly participate in heavy physical exercise"), BMI, age and sex without performing invasive fitness tests.

To determine whether regret plays a role in any potential effect, regret proneness is measured via the Regret Scale (RS; Schwartz et al., 2002) and anticipated regret of not exercising via adjusted questions of the Regret Experience Measure (REM; Creyer & Ross Jr, 1999). Exercising motivation is measured with the Behavioral Regulation in Exercise Scale (BREQ2; Markland & Tobin, 2004), and mental well-being with the Warwick-Edinburgh Mental Well-being Scale (WEMWBS; Tennant et al., 2007). See Table 2 for all measurements.

Statistical methods

Participants are the primary unit of inference and are nested within gyms. After 13, weeks a multi-level logistic regression analysis will be performed with the attained week-goals as dependent variable. The three trial-arms will be modelled as fixed effects with adjustment for pre-trial attendance and gyms as random intercepts. The multi-level model estimates the treatment effects after 13 weeks with respect to baseline attendance, while accounting for the clustered data pattern. After 26 weeks, another multi-level analysis will be performed. In this model, the dependent variable is the frequency of attained week-goals over the past 13 weeks, both measured at 13 and 26 weeks after the start of the study. These measurements are nested

within participants who are in turn nested within gyms. The three trial-arms and time are treated as fixed effects. Random intercepts will be added for both the participants and the gyms. A random slope for time on the participant level will also be included. This model aims to estimate the treatment effect over time while adjusting for the clustering of the measurements within participants and gyms. After 52 weeks, a similar model with the same goal is fitted, but the frequency of attained week-goals in the last 26 weeks is added.

Intra cluster effects

While there are 50 to 60 participants in each treatment arm, every arm contains no more than two clusters. Treatment effects may therefore correlate with intra cluster effects. We will account for this in our analyses in multiple ways. First, by controlling for baseline attendance, we capture a-priori differences in attendance-rates between gyms. Second, we will test the robustness of our models by conducting a sensitivity analysis. We will perform the analyses described above with every gym excluded once and compare the outcomes of each model with the models including all gyms.

Covariates

In their meta-analysis, Haff et al., (2015) determined that the behavioral effects in three independent regret-based lotteries were independent of gender, age, race and education (and only an income above $87.500 lowered the odds). Hence, we have no empirically based expectations on any confounding. Nonetheless, we will try to adjust in our analyses for age, sex, level of education and income. If, as expected, these covariates have no influence on any intervention effect, we will leave them out of the models.

Secondary analyses

After 13, 26 and 52 weeks, longitudinal logistic regression analysis will be performed on weekly goal attainment (no = 0, yes = 1), to model the probability of success over time, conditional on the treatment. As above, the measurements are nested within participants, within gyms and therefore random intercepts for participants and gyms will be added to the model, together with random slopes for time on the participant level.

To explore the treatment effect on absolute gym attendance, fitness, body composition, well-being and perceived health, multi-level linear regression models with random intercepts for participants and gyms, together with random slopes for time on the participant level will also be fitted.

To explore if participants in the lottery-arms anticipate more regret than participants in the control-arm do, a one-way ANOVA will be performed with anticipated regret as dependent variable and treatments as independent variable. We will also perform explorative analysis within the lottery-arms to determine whether regret-proneness and exercising motivation influences gym attendance.

Retention

As a compensation for their participation in the study, participants and High Five employees are promised and given three lottery tickets for the Dutch State Lottery (representing €45), regardless of gym attendance. Based on the principle of reciprocity (Cialdini, 2009), the tickets and online questionnaires are offered simultaneously to increase the chance of timely response by participants. To further prevent attrition and increase commitment to the study, we will hand all participants multiple gadgets with the university logo throughout the study (i.e., a water flask, duffle bag, towel, and mug) and send them a birthday card on behalf of the study staff.

Ethics

The study protocol, information letters and informed consent were reviewed and approved by the Tilburg University Ethical Committee. The study is registered in the Dutch Trial Register (NTR5559) and data storage is in accordance with guidelines offered by the data management department of the Institute for Public Health and the Environment and the Tilburg University quality guidelines for scientific research. All lottery drawings are performed by the independent Game Management department of the Dutch State Lottery (De Nederlandse Staatsloterij) under supervision of a notary. The State Lottery also formalized the lotteries' Terms and Conditions in accordance with the Dutch lottery legislation.

BASELINE CHARACTERISTICS

Table 3 displays the baseline characteristics of participants by study arm. As mentioned, based on the meta-analysis by Haff et al. (2015) we expect that any treatment effect will not vary with demographics. Furthermore, Zeelenberg & Pieters (2004) found that the influence of regret on decision-making is primarily influenced by expected counterfactual feedback and independent of attitudes and social norms. Therefore, aside from our eligibility criteria, there was no psychologically or demographically based recruitment strategy.

The gyms screened 182 candidates for eligibility, 163 of which were included in the study (69.3% male). Of the 163 participants, 48 were included in the control-arm, 60 in the short-term lottery-arm and 55 in the long-term lottery-arm. The majority of the participants is Dutch (86%) and has a monthly net income between €1000 and €3000 (72%). A total of four participants did not respond to the online survey. Weight, fat percentage and BMI were obtained in the first week of the trial. If participants did not attend their gym in the first week, weight at enrolment was used. Only self-reported weekly gym attendance is displayed, pending the data extraction from the gym registries.

Table 3. Participant Baseline characteristics per arm.

Characteristic	control-arm	short-term lotteries	long-term lottery
	(n = 48)	(n = 60)	(n = 55)
Age, mean (SD)	50 (9.84)	49.3(9.33)	45(9.58)
Gender, no. (%)			
Female	16 (33.3)	21 (35)	13 (23.6)
Male	32 (66.7)	39 (65)	42 (76.4)
No survey response, no. (%)	3 (6.25)	0 (0)	1 (1.82)
Nationality, no. (%)			
Dutch	36 (80)	52 (86.7)	52 (94.5)
Other	12 (20)	8 (13.3)	3 (5.5)
Education, no. (%)			
Pre-vocational education	3 (7.9)	7 (11.5)	4 (7.3)
Pre-university education	3 (6.7)	2 (3.3)	10 (18.2)
Senior vocational training	11 (24.4)	20 (33.3)	5 (9.1)
Vocational colleges	19 (42.2)	15 (25)	23 (41.8)
University education	9 (20)	15 (25)	10 (18.2)
Other	0 (0)	1 (1.7)	2 (3.6)
Monthly net income, no. (%)			
<€1000	0 (0)	0 (0)	1 (1.8)
€1000 to €2000	10 (20.8)	6 (10)	3 (5.5)
€2000 to €3000	19 (39.6)	32 (53.3)	24 (43.6)
€3000 to €4000	8 (16.7)	15 (25)	19 (34.5)
€4000 to €5000	2 (4.2)	1 (1.7)	2 (3.6)
€5000 tot €6000	0 (0)	2 (3.3)	1 (1.8)
> €6000	1 (2.1)	0 (0)	0 (0)
Did not wish to answer	5 (10.4)	4 (6.7)	4 (7.3)
Weekly gym attendance frequency * (self-report), mean (SD)	1.82 (0.88)	1.46 (1.17)	1.55 (1.04)
Weight, mean (SD)	90.14 (14.38)	96.12 (14.12)	96.6 (13.94)
Fat percentage, mean (SD)	33.78 (6.32)	35.52 (7.54)	36.83 (9.22)
BMI, mean (SD)	28.9 (3.20)	30.4 (3.73)	30.19 (3.47)
Obese, no.(%)	13 (27.1)	23 (38.3)	26 (47.3)
VO2max	31.05 (9.23)	28.13 (8.75)	33.63 (8.20)
Perceived Health (0-100)	71.53 (13.99)	68.18 (16.81)	68.48 (12.81)
Mental Well-being (WEMWBS, 1-5)	3.77 (0.55)	3.56 (0.60)	3.74 (0.61)
Regret Proneness (RS, 1-7)	3.41 (1.02)	3.45 (1.14)	3.20 (1.17)
Exercising motivation (BREQ2, 1-5)			
Amotivation	1.3 (0.47)	1.43 (0.52)	1.20 (0.48)
External Regulation	1.24 (0.51)	1.42 (0.79)	1.20 (0.37)
Introjected Regulation	2.26 (0.90)	2.22 (0.90)	2.17 (0.96)
Identified Regulation	3.99 (0.63)	3.58 (0.72)	4.01 (0.75)
Intrinsic Regulation	3.73 (0.92)	3.67 (0.94)	3.98 (0.94)

*Participants answered the question; "On average, how often per week did you attend the gym in the last two months?"

DISCUSSION

Above we described the design of Commitment Lotteries, a cluster randomized trial to test the effect of different lottery-based commitment devices on regular gym attendance. With this trial, we aim to 1) assist participants in attaining their exercising goals and 2) hereby gain practical and theoretical insights on effective commitment devices and 3) the possible health benefits.

Haisley et al. (2012), Kimmel et al. (2012), Volpp et al. (2008) and Patel et al., (2016) successfully performed similar lottery-trials to improve several health behaviors and found impressive behavior changes. The novelty of our trial lies in the setting, the behavioral outcome of interest (gym attendance), the long-term lottery deadline and the assessment of psychological constructs. Our findings can contribute to identifying behavioral economic methods to improve health behavior.

A limitation of our trial is that in order to obtain the lottery prize, participants are required to exercise at their gym and not at home or outside, which may be a threshold to engage in PA. A benefit of this approach is that we can reliably assess exercising frequency and that safe exercise is supervised by professionals. Another point of attention is that the majority (69.3%) of our sample is male, which is similar to the total population of the six gyms (69.8% male) and thus a consequence of research in the field. Although we expect not, we will examine if sex influences any results and we will discuss the practical implications of our findings.

The worldwide prevalence of overweight, obesity and associated health care costs have risen greatly over the past decades (Bray et al., 2016). Moreover, physical unfitness may be(come) an even greater risk on all-cause mortality (Barry et al., 2014). Given the benefits of exercise, the many intentions to lose weight and exercise (Boshove, 2014; Kooreman & Prast, 2010) and the well documented gap between intentions and behavior (DellaVigna & Malmendier, 2006; Sheeran, 2002), it is of vital importance to develop effective and readily applicable programs aimed at assisting individuals in attaining their exercise goals.

Ideally, this study will show if Commitment Lotteries yield long-term behavioral, physiological and psychological changes at low costs. Notably, the costs of health behavior change in this trial are relatively low because of the use of psychological theory in amplifying the perceived value of goal-attainment. That is, the expected value of goal attainment in any week is about €1.75, whereas the value of the anticipated regret in any week is €100.

One could argue that (government funded) rewarding of health behaviors may stress the public's sense of honesty and solidarity. It is then important to note that participants are not typically rewarded for their healthy behaviors because the majority of the exercising

participants receives no monetary prize. Furthermore, many commonly used commitment devices are self-funded and people voluntarily put their own money at stake to restrict their future choices and actions (John et al.,2011; Patel et al., 2016; Royer, Stehr, & Sydnor, 2012; Volpp et al., 2008). Hence, people looking for a form of commitment may also be willing to pay for their lottery-tickets in long-term applications.

Moreover, if insurers, governments, gym enterprises, human resource professionals or joint-forces of the preceding parties would assist the needing and willing population in exercising regularly at minimal costs, one could question why the lotteries should only serve as a short-term commitment device. Accordingly, one could argue in favor of a continuous application, where, if successful, the commitment device is part of integrated healthcare settings.

REFERENCES

Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: self- control by precommitment. *Psychological Science, 13*(3), 219-224.

Barry, V. W., Baruth, M., Beets, M. W., Durstine, J. L., Liu, J., & Blair, S. N. (2014). Fitness vs. fatness on all-cause mortality: a meta-analysis. *Progress in Cardiovascular Diseases, 56*(4), 382-390.

Biddle, S. J. H., & Fuchs, R. (2009). Exercise psychology: a view from Europe. *Psychology of Sport and Exercise, 10*(4), 410-419.

Boshove, D. (2014). *Goede voornemens 2015*. Customer Intelligence ING.

Bray, G. A., Fruhbeck, G., Ryan, D. H., & Wilding, J. P. (2016). Management of obesity.

The Lancet, 387(10031), 1947-1956.

Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment devices. *Annual Review of Economics, 2*(1), 671-698.

CBS, Central Bureau voor de Statistiek., & RIVM, Rijksinstituut voor Volksgezondheid en Milieu. (2017). *Gezondheidsenquête/Leefstijlmonitor*. Retrieved from statline.cbs.nl.

Cialdini, R. B. (2009). *Influence: Science and Practice*. Boston: Pearson Education.

Creyer, E. H., & Ross Jr, W. T. (1999). The development and use of a regret experience

measure to examine the effects of outcome feedback on regret and subsequent choice. *Marketing Letters, 10*(4), 373-386.

Dai, H., Milkman, K. L., & Riis, J. (2014). The fresh start effect: temporal landmarks motivate aspirational behavior. *Management Science, 60*(10), 2563-2582.

Dellavigna, S., & Malmendier, U. (2006). Paying not to go to the gym. *The American Economic Review, 96*(3), 694-719.

Fogelholm, M., & Kukkonen-Harjula, K. (2000). Does physical activity prevent weight

gain – a systematic review. *Obesity Reviews, 1*(2), 95-111.

Frijda, N. H. (2007). *The laws of emotion*. Mahwah: Lawrence Erlbaum Associates Publishers.

Given, B. A., Keilman, L. J., Collins, C., & Given, C. W. (1990). Strategies to minimize

attrition in longitudinal studies. *Nursing Research, 39*(3), 184-187.

Gulliford, M. C., Ukoumunne, O. C., & Chinn, S. (1999). Components of variance and

intraclass correlations for the design of community-based surveys and intervention studies: data from the health survey for England 1994. *American Journal of Epidemiology, 149*(9), 876-883.

Haff, N., Patel, M. S., Lim, R., Zhu, J., Troxel, A. B., Asch, D. A., & Volpp, K. G. (2015).

The role of behavioral economic incentive design and demographic characteristics in financial incentive-based approaches to changing health behaviors: a meta-analysis. *American Journal of Health Promotion, 29*(5), 314-323.

Haisley, E., Volpp, K. G., Pellathy, T., & Loewenstein, G. (2012). The impact of alternative incentive schemes on completion of health risk assessments. *American Journal of Health Promotion, 26*(3), 184-188.

Haskell, W. L., Lee, I. M., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A., . . .

Bauman, A. (2007). Physical activity and public health: updated recommendation for adults from the American college of sports medicine and the American heart association. *Medicine & Science in Sports & Exercise, 39*(8), 1423-1434.

Hassmén, P., Koivula, N., & Uutela, A. (2000). Physical exercise and psychological well- being: a population study in Finland. *Preventive Medicine, 30*(1), 17-25.

Hildebrandt, V.H., Ooijendijk, W.T.M., Hopman-Rock, M.(2007). *Trendrapport bewegen en gezondheid 2004/2005*. Hoofddorp/Leiden: TNO

Hulsegge, G., Van Der Schouw, Y. T., Daviglus, M. L., Smit, H. A., & Verschuren, W. M. (2016). Determinants of attaining and maintaining a low cardiovascular risk profile - the Doetinchem cohort study. *The European Journal of Public Health, 26*(1), 135-140.

Jackson, A. S., Blair, S. N., Mahar, M. T., Wier, L. T., Ross, R. M., & Stuteville, J. E. (1990). Prediction of functional aerobic capacity without exercise testing. *Medicine and Science in Sports and Exercise, 22*(6), 863-870.

John, L., Loewenstein, G., Troxel, A., Norton, L., Fassbender, J., & Volpp, K.G. (2011).

Financial incentives for extended weight loss: a randomized, controlled trial. *Journal of General Internal Medicine, 26*(6), 621-626.

Kettunen, O., Vuorimaa, T., & Vasankari, T. (2015). A 12-month exercise intervention decreased stress symptoms and increased mental resources among working

adults-results perceived after a 12-month follow-up. *International Journal of Occupational Medicine & Environmental Health, 28*(1), 157-168.

Kimmel, S. E., Troxel, A. B., Loewenstein, G., Brensinger, C. M., Jaskowiak, J., Doshi, J. A.,

. . . Volpp, K. (2012). Randomized trial of lottery-based incentives to improve warfarin adherence. *American Heart Journal, 164*(2), 268-274.

Kooreman, P., & Prast, H. (2010). What does behavioral economics mean for policy? challenges to savings and health policies in the Netherlands. *De Economist, 158*(2), 101-122.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics, 112*(2), 443-477.

Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet, 380* (9838), 219-229.

Markland, D., & Tobin, V. (2004). A modification to the behavioural regulation in exercise questionnaire to include an assessment of amotivation. *Journal of Sport and Exercise Psychology, 26*(2), 191-196.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review, 89* (1), 103-124.

Patel, M. S., Asch, D. A., Roy Rosin, M., Small, D. S., Bellamy, S. L., Heuer, J., . . . Volpp, K.G. (2016). Framing financial incentives to increase physical activity among overweight and obese adults. *Annals of Internal Medicine, 164*(6), 385-394.

Richard, R., De Vries, N. K., & Van Der Pligt, J. (1998). Anticipated regret and precautionary sexual behavior. *Journal of Applied Social Psychology, 28*(15), 1411-1428.

Rogers, T., Milkman, K. L., & Volpp, K. G. (2014). Commitment devices: using initiatives to change behavior. *JAMA, 311*(20), 2065-2066.

Royer, H., Stehr, M. F., & Sydnor, J. R. (2012). Incentives, commitments and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company. *American Economic Journal: Applied Economics, 7*(3), 51-84.

Schlam, T. R., Wilson, N. L., Shoda, Y., Mischel, W., & Ayduk, O. (2013). Preschoolers' delay of gratification predicts their body mass 30 years later. *The Journal of Pediatrics, 162*(1), 90-93.

Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: happiness is a matter of choice. *Journal of Personality and Social Psychology, 83*(5), 1178.

Sheeran, P. (2002). Intention—behavior relations: a conceptual and empirical review.

European Review of Social Psychology, 12(1), 1-36.

Sniehotta, F. F., Scholz, U., & Schwarzer, R. (2005). Bridging the intention–behaviour gap: planning, self-efficacy, and action control in the adoption and maintenance of physical exercise. *Psychology & Health, 20*(2), 143-160.

Stutzer, A., & Meier, A. N. (2015). Limited self-control, obesity, and the loss of happiness. *Health Economics*, 25(11), 1409-1424.

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., . . . Stewart-Brown, S. (2007). The Warwick-Edinburgh mental well-being scale (WEMWBS): development and UK validation. *Health and Quality of Life Outcomes, 5*(63).

Tsukayama, E., Toomey, S. L., Faith, M. S., & Duckworth, A. (2010). Self-control as a protective factor against overweight status in the transition from childhood to adolescence. *Archives of Pediatrics & Adolescent Medicine, 164*(7), 631-635.

Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial incentive–based approaches for weight loss: a randomized trial. *JAMA, 300*(22), 2631-2637.

Zeelenberg, M., Nelissen, R. M., Breugelmans, S. M., & Pieters, R. (2008). On emotion specificity in decision making: why feeling is for doing. *Judgment and Decision Making, 3*(1), 18-27.

Zeelenberg, M., & Pieters, R. (2004). Consequences of regret aversion in real life: the case of the Dutch postcode lottery. *Organizational Behavior and Human Decision Processes, 93*(2), 155-168

CHAPTER 4

Commitment lotteries promote physical activity among overweight adults: a cluster randomized trial

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Schipper, M., Zeelenberg, M., Berkhout, S., Polder, J.J., & Prast, H. M. (2018). Commitment lotteries promote physical activity among overweight adults- a cluster randomized trial. *Annals of Behavioral Medicine*, 52(4), 342-351.

ABSTRACT

The World Health Organization has identified physical inactivity as the fourth leading risk factor for global mortality. People often intend to engage in physical activity on a regular basis, but have trouble doing so. To realize their health goals, people can voluntarily accept deadlines with consequences that restrict undesired future behaviors (i.e., commitment devices). We examined if lottery-based deadlines that leverage regret aversion would help overweight individuals in attaining their goal of attending their gym twice per week. At each deadline a lottery winner was drawn from all participants. The winners were only eligible for their prize if they attained their gym-attendance goals. Importantly, nonattending lottery winners were informed about their forgone prize. The promise of this counterfactual feedback was designed to evoke anticipated regret and emphasize the deadlines. Six corporate gyms with a total of 163 overweight participants were randomized to one of three arms. We compared 1) weekly short-term lotteries for 13 weeks; 2) the same short-term lotteries in combination with an additional long-term lottery after 26 weeks; and 3) a control arm without lotteries.

After 13 weeks, participants in the lottery arms attained their attendance-goals more often than participants in the control arm. After 26 weeks, we observe a decline in goal attainment in the short-term lottery arm and the highest goal attainment in the long-term lottery arm. With novel applications, the current research adds to a growing body of research that demonstrates the effectiveness of commitment devices in closing the gap between health goals and behavior.

Physical activity (PA) is a key behavioral determinant of individual and public health (Barry et al., 2014; Haskell et al., 2007). Regular PA contributes to cardiovascular fitness and weight management, and reduces the risks of, among others, cardiovascular disease, cancers, diabetes mellitus type 2 and obesity (Fogelholm & Kukkonen-Harjula, 2000; Hulsegge et al., 2016; Lee et al., 2012). Consequently, the World Health Organization and governments worldwide recommend citizens to exercise on a regular basis (Haskell et al., 2007; Kahlmeier et al., 2015; WHO, 2010). Despite ample endorsements and many intentions to lose weight and exercise regularly (Baradel et al., 2009; Nicklas, Huskey, Davis, & Wee, 2012), 79% of Americans and 66% of Europeans do not meet recommended levels of PA (CDC, 2014; EC, 2014). Likewise, 74% of Americans and 62% of Europeans are overweight (Body Mass Index ≥ 25; Flegal, Carroll, Kit, & Ogden, 2012; WHO, 2015).

Although people often intend to change their behavior and engage in PA on a regular basis, they systematically fail to do so (DellaVigna & Malmendier, 2006; Rhodes & Dickau, 2012). Behavioral economics, operating at the intersection of economics and psychology (Bickel, Moody, & Higgins, 2016), provides insights that help explain the difficulties of behavior change, including present bias: the human tendency to disproportionally overweigh costs and benefits that are immediate over those that are delayed (Laibson, 1997; O'Donoghue & Rabin, 1999; Strotz, 1955). Correspondingly, long-term health-goals are widely adopted, but are mostly not fully achieved (Acland & Levy, 2015; Baradel et al., 2009): despite previous intentions, the immediate costs (e.g. exercising) overshadow the delayed benefits (e.g. good health), resulting in procrastination (Ariely & Wertenbroch, 2002).

To not fall prey to this pattern, people can voluntarily accept meaningful deadlines that impose potential costs on undesired future behaviors, known as commitment devices (Ariely & Wertenbroch, 2002; Rogers, Milkman, & Volpp, 2014). A common application of a commitment device is the 'deposit contract', via which individuals voluntary deposit money that they will lose if they fail to achieve a predetermined personal goal at a deadline (Giné, Karlan, & Zinman, 2010; John et al., 2011; Rogers et al., 2014). By restricting behavior ahead of time, commitment devices strategically avert present biased tendencies and can hereby help individuals in conserving their intended exercising behavior (Bryan, Karlan, & Nelson, 2010).

Although physical inactivity is hazardous in all BMI-ranges (Barry et al., 2014; Mainous, Tanner, Anton, Jo, & Luetke, 2017), overweight (BMI ≥ 25) and obese (BMI ≥ 30) individuals are especially likely to benefit from commitment devices for PA because they generally exercise less than normal-weight individuals (CBS & RIVM, 2017), while regular PA can contribute to weight- loss and management. Besides, overweight and obesity have been associated with a relatively strong disposition to overweigh the present over the future (i.e. present bias; Bickel et al., 2014; Ikeda, Kang, & Ohtake, 2010; Schlam, Wilson, Shoda, Mischel, & Ayduk, 2013; Weller, Cook, Avsar, & Cox, 2008) and commitment devices are designed to preempt this.

Drawing on previous applications of behavioral economics in supporting health behavior change (see: Haff et al., 2015), we tested multiple lottery-deadlines intended to help overweight adults in attaining their gym-attendance goals. Research suggests that people are generally regret averse, meaning that they anticipate regret and often make decisions that minimize regret in the future (Zeelenberg & Pieters, 2007). The lottery-deadlines were designed to leverage regret aversion by incorporating a key feature of the Dutch postal code lottery (2.5 million players per drawing). In the postal code lottery all postal codes can win, but only the residents who purchased tickets get a prize. Inevitably, residents of the winning region who did not purchase tickets discover that they would have had a prize if they had decided differently in the past. Accordingly, regret aversion has been found to motivate the decision to purchase lottery tickets (Zeelenberg & Pieters, 2004).

In the present study, participants committed to their goal of attending their gym twice per week by voluntarily accepting multiple lottery-deadlines. At each lottery-deadline a winner was drawn from all participants. The winners, however, were only eligible to receive their prize if they attained their gym-attendance goals. Importantly, lottery winners who did not attain their goal were informed about their forgone prize. The promise of feedback on 'what would have been' was designed to emphasize the possibility of regret at the deadlines (Zeelenberg, 1999).

We set up a three-arm cluster randomized trial across six gyms to examine if commitment lotteries would support overweight adults in attaining their goal of attending their gym twice per week. We compared 1) weekly short-term lotteries for 13 weeks; 2) the same short-term lotteries in combination with an additional long-term lottery after 26 weeks; and 3) a control arm without lotteries. We examined the effect of the lottery interventions on weekly individual goal-attainment over 13, 26 and 52 weeks compared to a control arm. This article reports on the results after 13 and 26 weeks.

We hypothesized that after 13 weeks, participants in both lottery-arms would be more likely to attain their week-goals than participants in the control-arm. Behavioral economic commitment schemes generally result in behavior change in the short-run, but the changes are mostly not fully maintained (Patel, 2016; Royer, Stehr, & Sydnor, 2012; Volpp et al., 2008). Therefore, we expanded the short-term deadlines with an additional long-term deadline to test if this would promote long-term goal attainment. Hence, after 26 weeks, we expected a decline in goal attainment in the short-term lottery-arm and the highest goal attainment in the long-term lottery arm (Van der Swaluw et al., 2016).

METHOD

Design

The rationale and protocol of this trial have been published elsewhere (Van der Swaluw et al., 2016). The design is a three-arm, parallel group, cluster randomized trial running for 52 weeks with 163 participants in six corporate gyms (clusters) across the Netherlands. Figure 1 displays the study design and flow. The trial protocol and materials were reviewed and approved by the Tilburg University Ethical Review Board (EC-2014.42a). The study is registered in the Dutch Trial Register (NTR5559) and lottery drawings were performed by the independent Game Management Department of the Dutch State Lottery under supervision of a notary.

Participants & Enrollment

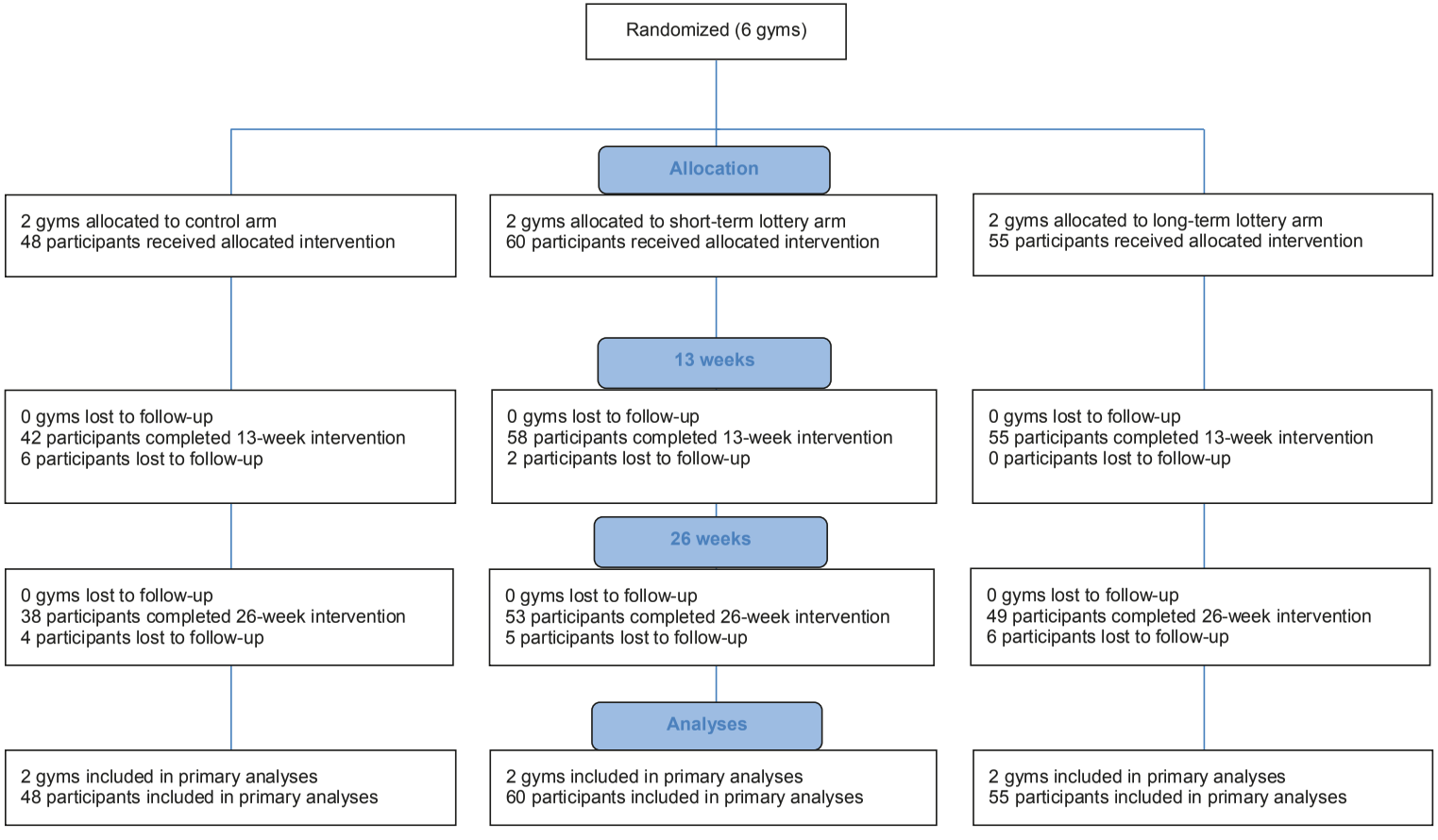
Gyms were eligible to participate if the managers expressed their interest in scientific research prior to randomization. The six gyms were a randomly selected from a convenience sample from 36 corporate gym-sites across the Netherlands, hosted by fitness enterprise High Five. Next to written information and an oral briefing, gyms received a tailored video containing the rationale and protocol of the trial. With a standardized recruiting text, provided to the gyms, gym managers recruited new and existing members who were looking for a commitment device for regular exercise, via email, company webpages and in person. The material summarized the nature and procedure of the study and directed candidates to the gym personnel. We aimed to recruit a minimum of 25 participants per gym, but allowed gyms to screen more participants.

Candidates were eligible if they explicitly stated to have the goal to exercise twice or more per week, were overweight ($25 \leq BMI < 40$), between the age of 18–65 and had not planned a leave of absence of more than 4 weeks in the first 26 weeks of the trial. Together with the gym-personnel, candidates weighed on a provided scale (KERN™; 0.1% precision) and filled out a digital questionnaire which immediately identified whether the candidate was eligible or not. After providing informed consent, applicants were entered into the study.

Interventions

This trial compares two intervention-arms to one control-arm. The interventions pertain to the participant level. The American College of Sports Medicine and the American Heart Association endorse vigorous exercise for 20 min, three days a week, and muscular strength and endurance training two days a week (Haskell et al., 2007). Consequently, setting the goal of attending the gym two days a week was considered beneficial, while challenging but attainable (Van der Swaluw et al., 2016). Therefore, participants in all three arms set the goal to attend their gym twice per week (the week-goal) and were handed a randomly generated three-digit study-ID prior to the start of the trial. Upon entering their gym, all participants were required to register their attendance with their study-ID on trial-iPads, provided to the gyms. All participants were offered a monthly overview of their attendance via email.

Figure 1: Study design and flow of gyms and participants.



Intervention arm 1: short-term lottery.

For 13 weeks, participants in this arm participated in a free weekly lottery worth €100 each drawing. The winning number (study-ID) was drawn from all participants in this arm (participants knew that they could always win the lottery) and communicated to all via text message and email (participants knew that they would always learn the outcome). The winners only received their prize if they attended their gym at least twice that week (the week-goal). Importantly, lottery winners who did not attain their week-goal were informed about their forgone prize. All other participants knew whether the week-prize was awarded or not, but not to whom. Notably, every new week offered a new opportunity to win and to keep attaining exercise goals, regardless of prior success. This feature facilitates the human inclination to use temporal landmarks (e.g. Mondays) as a fresh start by relegating misfortune to the past (Dai, Milkman, & Riis, 2014). The weekly expected monetary value for a fully compliant subject was 1/60 = €1.67. Note, however, that the lotteries were designed to emphasize the deadlines and not as a payment.

Intervention arm 2: long-term lottery.

The intervention in this arm was identical to the short-term lottery arm in the first 13 weeks. The weekly expected monetary value for a fully compliant subject was 1/55 = €1.82. Additionally, weeks 14-26 were also part of the intervention (participants knew this prior to the start of the trial). After week 26, a luxury vacation-cheque for the winner and four friends or family members (communicated as such to participants, worth €5400) was awarded. The winning number was again drawn from all participants and communicated to all via text message and email. Participants were informed that the prize could only be claimed if the winner would attain the week-goal in at least 9 of the second 13 weeks (70% between weeks 14-26). Because weeks 14-26 fell in the national holiday season, the 9:13 success ratio provided participants the opportunity to enjoy a vacation and still be eligible for their prize. Participants knew that if the winner would not meet the requirements for obtaining the prize, he or she would receive a small consolation prize and another number would be drawn until the prize could be claimed.

Control arm.

In the control arm, participants also set the goal to attend the gym twice per week and were monitored in their attendance and secondary outcomes, but no commitment devices were offered. As such, the lotteries were the only designed differences between control- and intervention arms. Participants in the control arm were also offered monthly statistics on their performance via email.

Outcomes and measures

The primary outcome of interest was goal-attainment (week-gym attendance $\geq$ 2) measured at the participant level and assessed by requiring participants to check in to the trial-iPad when entering their gym. Baseline attendance levels, nationality, age, sex, education and income level were assessed via questionnaires and are displayed in Table 1.

Sample size and randomization

The sample size calculation for this trial has been reported in detail before (Van der Swaluw et al., 2016). Anticipating a 0.35 difference between proportions, based on meta-analysis by Haff et al. (2015), and accounting for the clustered design, we estimated a required sample size of 36 per arm and aimed to include at least 50 participants per arm, allowing for 25-30% attrition. No within-gym randomization was performed to avoid intervention contamination, maintain blinding at the participant level and to minimize the administrative burden for the gym personnel. Therefore, every trial-arm included two gyms. Participants were informed that there were two gyms in their arm, but not about the content of the interventions in the other gyms and arms. Based on anonymized member data, we were able to distinguish three gyms with a relatively high- and three gyms with a relatively low proportion of overweight members. By computer generation, first high-proportion gyms and next low-proportion gyms were randomly allocated to one of three arms, preventing large differences in enrollment time.

Table 1. Baseline participant characteristics displayed by study arm.

Characteristic	Control (n = 48)	Short-term lotteries (n = 60)	Long-term lottery (n = 55)
Age, mean (SD)	50 (9.84)	49.3(9.33)	45(9.58)
Gender, no. (%)			
Female	16 (33.3)	21 (35)	13 (23.6)
Male	32 (66.7)	39 (65)	42 (76.4)
No survey response, no. (%)	3 (6.25)	0 (0)	1 (1.82)
Nationality, no. (%)			
Dutch	36 (80)	52 (86.7)	52 (94.5)
Other	12 (20)	8 (13.3)	3 (5.5)
Education, no. (%)			
Pre-vocational education	3 (7.9)	7 (11.5)	4 (7.3)
Pre-university education	3 (6.7)	2 (3.3)	10 (18.2)
Senior vocational training	11 (24.4)	20 (33.3)	5 (9.1)
Vocational colleges	19 (42.2)	15 (25)	23 (41.8)
University education	9 (20)	15 (25)	10 (18.2)
Other	0 (0)	1 (1.7)	2 (3.6)
Monthly net income, no. (%)			
<€1000	0 (0)	0 (0)	1 (1.8)
€1000 to €2000	10 (20.8)	6 (10)	3 (5.5)
€2000 to €3000	19 (39.6)	32 (53.3)	24 (43.6)
€3000 to €4000	8 (16.7)	15 (25)	19 (34.5)
€4000 to €5000	2 (4.2)	1 (1.7)	2 (3.6)
€5000 tot €6000	0 (0)	2 (3.3)	1 (1.8)
> €6000	1 (2.1)	0 (0)	0 (0)
Did not wish to answer	5 (10.4)	4 (6.7)	4 (7.3)
Baseline gym attendance*, mean (SD)	1.82 (0.88)	1.46 (1.17)	1.55 (1.04)
Weight, mean (SD)	90.14 (14.38)	96.12 (14.12)	96.6 (13.94)
Fat percentage, mean (SD)	33.78 (6.32)	35.52 (7.54)	36.83 (9.22)
BMI, mean (SD)	28.9 (3.20)	30.4 (3.73)	30.19 (3.47)
Obese, no.(%)	13 (27.1)	23 (38.3)	26 (47.3)

*Participants answered the question; "On average, how often per week did you attend the gym in the last two months?"

Statistical methods

Analyses followed the intention-to-treat principle and were conducted in R version 3.3.1 and SPSS Statistics version 22 (IBM Corp, Armonk, NY) with statistical significance set at $p < 0.05$. Goal attainment was evaluated binary (0 = no, 1 = yes) at the participant level. Multivariate logistic mixed models were used to assess between-arm differences in goal-attainment between weeks 1-13 and weeks 14-26 controlled for baseline PA, age and sex. The control arm was modeled as the reference category and gyms were modeled as random intercepts.

In the mixed models, intervention effects are adjusted for the dependence of the outcome within gyms and adjusted for baseline PA differences. As such, in estimating the coefficients, the mixed models account for the clustered data pattern. To further inspect within-gym effects, we additionally performed sensitivity analyses by excluding each gym from the models once and comparing effects from these models to effects in the complete model.

RESULTS

Table 2 displays the average frequency of goal attainment per 13 weeks. Additionally, Figure 2 displays the adjusted probabilities of goal attainment between weeks 1-13 and 14-26 per arm.

Table 2. Average frequency of successful weeks (gym attendance ≥2) per study period.

	Weeks 1 - 13	**Weeks 14 - 26**	**Weeks 1 - 26**
	Mean (SD)	Mean (SD)	Mean (SD)
Study arm			
Control	3.54 (4.03)	3.38 (4.06)	6.92 (7.45)
Short-term lotteries	7.33 (3.58)	3.18 (3.37)	10.52 (6.20)
Long-term lottery	8.31 (4.05)	6.25 (4.38)	14.52 (7.84)

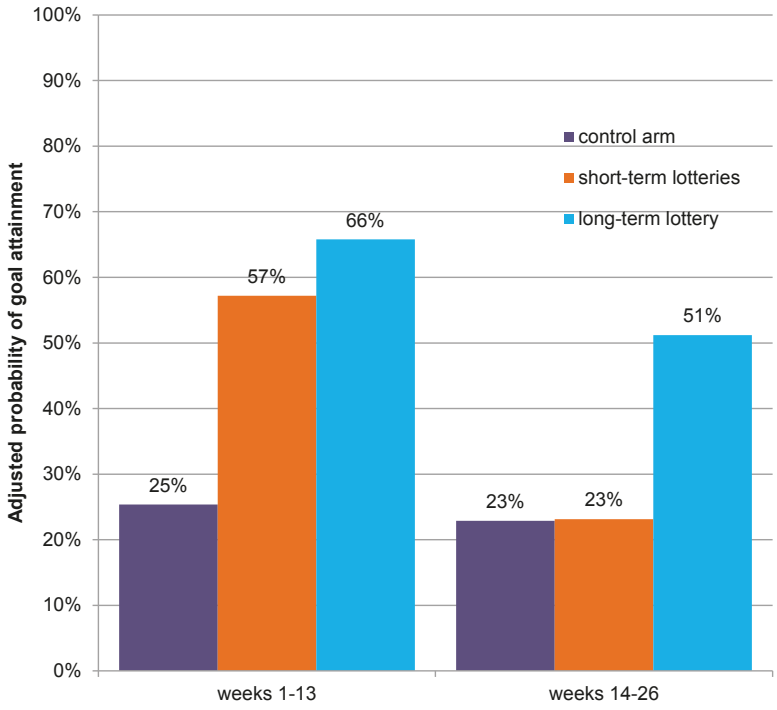
Weeks 1-13

In both lottery arms, 8 of the 13 lottery winners (62%) received their prize. Participants in both lottery arms were more likely to attain their week-goal than participants in the control arm. On average, participants in the control arm attained 27% of their week-goals opposed to 55% and 63% in the short-term lottery arm and long-term lottery arm respectively. Accordingly, the mixed logistic model (Table 3) showed a statistically significant intervention effect on goal attainment for the short-term lottery arm (OR= 3.39, 95% CI, 1.20 – 12.92) and the long-term lottery-arm (OR = 5.66, 95% CI, 1.72 – 18.66). The intervention effect did not differ significantly between both intervention arms (OR= 1.44, 95% CI, 0.44 – 4.70). The results of the sensitivity analyses were qualitatively similar to those based on primary analysis: the direction of effects in the sensitivity analyses did not diverge from the intervention effects in the complete model.

Weeks 14-26

On average, participants in the control arm and short-term lottery arm attained 25% and 24% of their week-goals respectively, whereas participants in the long-term lottery arm on average attained 48% of their week-goals. Participants were eligible to receive the long-term lottery if they attained their goal in at least 9 of the second 13 weeks. In total, 55% of participants in the long-term lottery arm attained the week-goal in ≥ 9 weeks. The mixed logistic model showed a statistically significant intervention effect on goal attainment for the long-term lottery (OR = 3.53, 95% CI, 1.28 – 9.77). Besides, participants in the long-term lottery arm were significantly more likely to attain their goals than participants in the short-term lottery arm (OR = 3.48, 95% CI, 1.27 – 9.57). In contrast to weeks 1-13, the likelihood of goal-attainment in the short-term lottery arm no longer differed significantly from the control arm (OR = 1.01, 95% CI, 0.37 – 2.80).

Figure 2. Adjusted probabilities of goal-attainment (week attendance ≥2)* between weeks 1-13 and 14-26, displayed by trial period.



*Adjusted for within-gym clustering, baseline attendance, age and sex.

The sensitivity analyses showed qualitatively similar intervention effects for the long-term lottery arm. The estimated coefficient of the short-term lottery arm was sensitive to exclusion of gyms from the control arm. The non-effect in the complete model became a negative effect when excluding the least performing gym in the control arm from the analyses. The effect became positive when excluding the best performing control-gym from the analyses.

Table 3. Logistic mixed models predicting goal attainment (week attendance ≥2) between weeks 1-13 and 14-26.

	Weeks 1 – 13 Odds Ratio (95% CI)	**Weeks 14-26** Odds Ratio (95% CI)
Study arm		
Short-term lotteries	3.93* (1.20-12.92)	1.01 (0.37-2.80)
Long-term lottery	5.66** (1.72-18.66)	3.53* (1.28-9.77)
Participant characteristics		
Baseline attendance	1.28** (1.17-1.41)	1.40** (1.27-1.55)
Age	1.00 (0.99-1.01)	1.02** (1.01-1.03)
Male vs. Female	0.54** (0.43-0.68)	0.73** (0.58-0.93)

* significant at $p < .05$
** significant at $p < .01$
Intra Cluster Correlation (weeks 1-13): 0.10, (weeks 14-26): 0.07

DISCUSSION

The results from this cluster randomized trial show that commitment lotteries can help overweight adults in attaining their goal of attending their gym twice per week. Participants who voluntarily committed to 13 weekly lottery deadlines were more likely to attain their goal of attending their gym twice per week than participants in the control arm. Furthermore, participants who were assigned to an additional lottery deadline after 26 weeks were more likely to attend their gym twice per week until 26 weeks than participants without this long-term lottery deadline.

Although this trial showed that weekly lotteries were effective in providing short-term commitment, goal-attainment decreased in absence of an additional long-term deadline. As expected, the additional long-term lottery deadline partly averted the decline in PA after an initial period of success.

The present findings expand knowledge on the use of commitment devices to facilitate behavior change. The effectiveness of- and demand for commitment devices has been illustrated in an increasing body of behavioral research. For example; people voluntarily restrict future spending (Beshears et al., 2015; Thaler & Benartzi, 2004), eating (Wertenbroch, 1998), or smoking (Giné

et al., 2010) to facilitate (retirement) saving, weight loss and quitting attempts. The present trial contributes with a novel behavioral context (gym attendance) and the application of a long-term lottery deadline.

To overcome present biased decision-making and procrastination, behavioral research generally recommends increasing immediate (costs) benefits of (un)desirable behaviors as a strategy for behavior change (Loewenstein, Brennan, & Volpp, 2007; Soman et al., 2005). In this reasoning, the effectiveness of the weekly lottery deadlines can be explained by their ability to impose nearby consequences on procrastination. A nearby deadline with the chance to win, but miss out on €100 limits the time window for action and hereby prioritizes the desired behavior. Previous studies have used comparable strategies to effectively support medication adherence (Kimmel et al., 2012), weight loss (Volpp et al., 2008) and walking (Patel, 2016).

In contrast to multiple nearby deadlines, a distant deadline interferes less with present biased preferences and leaves more time for procrastination. This was demonstrated in research by Ariely and Wertenbroch (2002) in which students' academic performance decreased when they accepted one distant deadline opposed to multiple nearby deadlines. However, in the present trial, the long-term lottery deadline partly averted the decline in goal-attainment that we observed in the short-term lottery arm after removal of the weekly lottery deadlines. The threat of learning that; "I would have had a free family vacation if I had decided differently in the past" (i.e. regret aversion) could be an explanation for this.

Regret in the future has the ability to influence health behaviors in the present by emphasizing the future consequences of current decisions (Chapman & Coups, 2006; Richard, De Vries, & Van der Pligt, 1998). Results from meta-analysis by Brewer, DeFrank, and Gilkey (2016) additionally show that the effect of anticipated inaction-regret (e.g., not exercising) on health behavior is unaffected by the temporal distance of the negative consequence. Therefore, in contexts where possible regret at a distant deadline is made salient, distant deadlines may avert present-biased decision-making similarly to multiple nearby deadlines. More research on deadline distance in relation to regret would valuably contribute to the open question of the optimal duration and interval of commitment devices (Rogers et al., 2014).

Scholars reviewing the effectiveness of commitment devices have concluded that the development of commitment devices is still in its early stages (Brocas, Carrillo, & Dewatripont, 2004; Bryan et al., 2010; Rogers et al., 2014). Although their design and acceptance have received considerable attention (Beshears et al., 2015; Laibson, 2015), it remains difficult to project which contextual and behavioral features optimize its uptake and cost-effectiveness (Halpern, Asch, & Volpp, 2012). Notably, the weekly lotteries and an additional long-term lottery were effective at only about €5 per participant per week (prizes ÷ participants ÷ weeks). Because previous research has demonstrated that people are willing to put their own money at stake

(Bryan et al., 2010; Royer et al., 2012) or pay premiums to restrict their future choices (Beshears et al., 2015), it would be valuable to explore if and when people would also be willing to pay for lottery tickets as a commitment to their health goals.

Evidently, the costs per participant decrease if the lotteries are accepted on a larger scale. To help understand the feasibility of voluntary commitment, O'Donoghue and Rabin (1999) have formalized the intuitive distinction between two extreme types of people: those who are fully aware about their future self-control difficulties (sophisticates) and those who are fully unaware (naïfs). Although both types of people may benefit from commitment devices, sophisticates are most likely to accept and profit from imposed deadlines (Acland & Levy, 2015; Bryan et al., 2010). It remains unclear, however 1) if commitment devices (or meaningful deadlines) are effective if 'sophisticates' accept commitment, but nonetheless have low intrinsic motivation to perform the targeted behavior and 2) how the acceptance and use of commitment devices with a financial component may ultimately affect intrinsic motivation. Answering these open questions would valuably contribute to the effectiveness and attractiveness of commitment lotteries. Further research on the feasibility of commitment devices should focus on these unresolved questions.

Despite the financial component of the present interventions, we designed and communicated the commitment lotteries as commitment devices rather than financial incentives. Commitment lotteries differ from traditional incentives in multiple ways. First, they differ in the problem that they target. Commitment devices aim to assist people who are initially motivated to exercise on a regular basis, but believe they will probably fail to do so without proper commitment.
In contrast, a financial incentive in its most traditional (neoclassical economic) sense is aimed at encouraging the unmotivated to become motivated due to the payment (Gneezy & Rustichini, 2000; Mantzari et al., 2015; Marteau, 2009). An incentive is thus a conditional cash transfer in order to increase the attractiveness of a certain behavior.
In a commitment lottery, the majority of participants received no payment (approximately 84%). Besides, the expected monetary value of weekly goal attainment was low (e.g. only about €1.73 in the first 13 weeks), which is substantially lower than traditional incentives (i.e., payments) for health behavior change (Charness & Gneezy, 2009; Mantzari et al., 2015; Rohde & Verbeke, 2017; Royer et al., 2012).

Second, financial incentives differ from commitment lotteries in their contingency. In order to be eligible for a traditional cash payment, one has to perform the targeted behavior. This does not exclude a variable payment (e.g. a lottery), but traditionally, lottery *participation* is the reward for a specific health behavior (Mantzari et al., 2015). In a commitment lottery, however, the imposed deadline is emphasized by the fact that all participants are included and can win, irrespective of their success. Therefore, commitment lotteries were not designed and communicated to participants as payments for attending the gym, but as a way to commit to an individual goal.

Although there are multiple essential differences between a commitment lottery and a traditional lottery or simple payment, commitment lotteries also hold a clear financial component that should not be disregarded as a factor influencing the present results. For this reason, it would be valuable to explore optimal prize sizes and willingness to pay for commitment lottery tickets (hereby attenuating the financial component).

A limitation of this trial is that, although 163 participants enrolled, only six units (gyms) were randomized. Randomization at the gym level, however, avoided intervention contamination within gyms and adapted best to daily practice. That being; scientific research is not the core business of gym enterprises and researchers are safest in assuming that it has low priority in daily practice. Therefore, gyms were likely to benefit from one intervention at a time. Future research in gym contexts could extend the number of gyms.
Another limitation is that we did not directly observe physical activity in the gyms and assumed that participants attended their gym to exercise. An interesting topic for forthcoming research could be the effect of commitment devices on changes in the duration of gym visits or improvements in exercising routine.
Not surprisingly, sensitivity analyses showed that between-gym variation in goal attainment was highest in the control arm: in absence of a homogeneous intervention, other, non-identified factors are likely to have had more influence on goal attainment. As a result, the non-difference between the short-term arm and the control arm between weeks 14-26 showed to be sensitive to exclusion of control gyms. Nonetheless, the most stringent interpretation of all results remains that the short-term lotteries are effective as long as they are present. An additional long-term deadline after weekly short-term deadlines was effective in partly preventing the decline in goal attainment after removal of the weekly deadlines.

The novel application of commitment lotteries to gym attendance has multiple benefits. First, health professionals recommend strength- and endurance training two days a week (Haskell et al., 2007), while gyms are principally equipped for this purpose. Second, offering commitment devices for gym attendance aligns with societal preferences: exercise in gyms is currently one of the most popular modes of exercise (Tiessen-Raaporst, 2015). Third, reliability of PA monitoring increases as participants can only register their exercise at the gym sites. Hence, gym contexts are well suited for testing commitment lotteries, while safe and suited exercise is supervised by trained professionals (Thompson, 2015).

Non-communicable diseases are responsible for approximately 70% of deaths worldwide and next to significantly affecting health and well-being (WHO, 2011), impose a substantial economic burden (Bloom et al., 2012). Given the significant role of modifiable behavior (e.g., exercising) in preventing non-communicable diseases and the increasing pressure on public health expenses (Bloom et al., 2012; Hoeymans et al., 2014), there is a need for innovative low cost approaches to health behavior change. The effectiveness of the use of personal emotions

and use of social contexts (Wally & Cameron, 2017) to support health behavior change shows promising directions in levering the impact of investments. Besides, it is not difficult to imagine possibilities for applying and further developing commitment lotteries in field settings. For example, innovative employers, governments, insurers, gyms, clinical health centers or consortia of such could offer commitment lotteries as a part of integrated care settings. In this manner, continuous supply and reminders of voluntary deadlines for health behavior change might help avert the return to old unwanted habits (Kaushal, Rhodes, Spence, & Meldrum, 2017).

CONCLUSION

Many people aim to exercise on a regular basis but fail to do so. Commitment lotteries were effective in supporting regular exercise and only as long as the threat of missing out on the lottery prize was present. Weekly short-term lotteries supported regular PA for 13 weeks and an additional long-term lottery after 26 weeks showed to partly avert the decline in goal-attainment after the 13 weekly lotteries. With novel applications, the current research adds to a growing body of research that shows the effectiveness of commitment devices in closing the gap between health goals and behavior.

REFERENCES

Acland, D., & Levy, M. R. (2015). Naiveté, projection bias, and habit formation in gym

Attendance. *Management Science, 61*(1), 146-160.

Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: self-control by precommitment. *Psychological Science, 13*(3), 219-224.

Baradel, L. A., Gillespie, C., Kicklighter, J. R., Doucette, M. M., Penumetcha, M., & Blanck, H. M. (2009). Temporal changes in trying to lose weight and recommended weight-loss strategies among overweight and obese Americans, 1996–2003. *Preventive Medicine, 49*(2), 158-164.

Barry, V. W., Baruth, M., Beets, M. W., Durstine, J. L., Liu, J., & Blair, S. N. (2014). Fitness vs. fatness on all-cause mortality: a meta-analysis. *Progress in Cardiovascular Diseases, 56*(4), 382-390.

Beshears, J., Choi, J. J., Harris, C., Laibson, D., Madrian, B. C., & Sakong, J. (2015). Self control and commitment: can decreasing the liquidity of a savings account increase deposits? *National Bureau of Economic Research, w21474*, 1–53.

Bickel, W. K., George Wilson, A., Franck, C. T., Terry Mueller, E., Jarmolowicz, D. P., Koffarnus, M. N., & Fede, S. J. (2014). Using crowdsourcing to compare temporal, social temporal, and probability discounting among obese and non-obese individuals. *Appetite, 75*, 82-89.

Bickel, W. K., Moody, L., & Higgins, S. T. (2016). Some current dimensions of the behavioral economics of health-related behavior change. *Preventive Medicine, 92* (C), 16-23.

Bloom, D. E., Cafiero, E., Jané-Llopis, E., Abrahams-Gessel, S., Bloom, L. R., Fathima, S., . . . Mowafi, M. (2012). *The global economic burden of noncommunicable diseases.* Geneva: World Economic Forum.

Brewer, N. T., DeFrank, J. T., & Gilkey, M. B. (2016). Anticipated regret and health behavior: A meta-analysis. *Health Psychology, 35*(11), 1264-1275.

Brocas, I., Carrillo, J. D., & Dewatripont, M. (2004). Commitment devices under self-control problems: An overview. In I. Brocas, & J. D. Carrillo (Eds.), *The Psychology of Economic Decisions, 2*, (pp. 49-67). Oxford: Oxford University Press.

Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment devices. *Annual Review of Economics, 2*(1), 671-698.

CBS, Central Bureau voor de Statistiek., & RIVM, Rijksinstituut voor Volksgezondheid en Milieu. (2017). *Gezondheidsenquête/Leefstijlmonitor.* Retrieved from statline.cbs.nl.

Centers for Disease Control and Prevention, & Division of Nutrition, P. A., and Obesity. (2014). *Facts about Physical Activity.* Retrieved from www.cdc.gov/physicalactivity/data/facts.htm

Chapman, G. B., & Coups, E. J. (2006). Emotions and preventive health behavior: worry, regret, and influenza vaccination. *Health Psychology, 25*(1), 82-90.

Charness, G., & Gneezy, U. (2009). Incentives to exercise. *Econometrica, 77*(3), 909-931.

Dai, H., Milkman, K. L., & Riis, J. (2014). The fresh start effect: Temporal landmarks motivate aspirational behavior. *Management Science, 60*(10), 2563-2582.

DellaVigna, S., & Malmendier, U. (2006). Paying not to go to the gym. *The American Economic Review*, 694-719.

European Commission: Strategy, Corporate Communication Actions and Eurobarometer Unit

(2014). Sport and physical activity. *Special Eurobarometer 412.*

Flegal, K. M., Carroll, M. D., Kit, B. K., & Ogden, C. L. (2012). Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. *JAMA, 307*(5), 491-497.

Fogelholm, M., & Kukkonen-Harjula, K. (2000). Does physical activity prevent weight gain – a systematic review. *Obesity Reviews, 1*(2), 95-111.

Giné, X., Karlan, D., & Zinman, J. (2010). Put your money where your butt is: a commitment contract for smoking cessation. *American Economic Journal: Applied Economics, 2* (4), 213-235.

Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all. *Quarterly Journal of Economics, 115*(3), 791-810.

Haff, N., Patel, M. S., Lim, R., Zhu, J., Troxel, A. B., Asch, D. A., & Volpp, K. G. (2015). The role of behavioral economic incentive design and demographic characteristics in financial incentive-based approaches to changing health behaviors: a meta-analysis. *American Journal of Health Promotion, 29*(5), 314-323.

Halpern, S. D., Asch, D. A., & Volpp, K. G. (2012). Commitment contracts as a way to health. *BMJ: British Medical Journal, 344*(1), e522-e522.

Haskell, W. L., Lee, I. M., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A., . . . Bauman, A. (2007). Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Medicine and Science in Sports and Exercise, 39*(8), 1423-1434.

Hoeymans, N., Van Loon, A., Van den Berg, M., Harbers, M., Hilderink, H., Van Oers, J., & Schoemaker, C. (2014). *Een gezonder Nederland: Volksgezondheid Toekomst Verkenning 2014*. Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu.

Hulsegge, G., van der Schouw, Y. T., Daviglus, M. L., Smit, H. A., & Verschuren, W. M. (2016). Determinants of attaining and maintaining a low cardiovascular risk profile-the Doetinchem Cohort Study. *European Journal of Public Health, 26*(1), 135-140.

Ikeda, S., Kang, M.-I., & Ohtake, F. (2010). Hyperbolic discounting, the sign effect, and the body mass index. *Journal of Health Economics, 29*(2), 268-284.

John, L., Loewenstein, G., Troxel, A., Norton, L., Fassbender, J., & Volpp, K. (2011). Financial incentives for extended weight loss: a randomized, controlled trial. *Journal of General Internal Medicine, 26*(6), 621-626.

Kahlmeier, S., Wijnhoven, T. M., Alpiger, P., Schweizer, C., Breda, J., & Martin, B. W. (2015). National physical activity recommendations: systematic overview and analysis of the situation in European countries. *BMC Public Health, 15*(1), 133-147.

Kaushal, N., Rhodes, R. E., Spence, J. C., & Meldrum, J. T. (2017). Increasing physical activity through principles of habit formation in new gym members: a randomized controlled trial. *Annals of Behavioral Medicine, 51*(4), 578-586.

Kimmel, S. E., Troxel, A. B., Loewenstein, G., Brensinger, C. M., Jaskowiak, J., Doshi, J. A., . . . Volpp, K. (2012). Randomized trial of lottery-based incentives to improve warfarin adherence. *American Heart Journal, 164*(2), 268-274.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics, 112*(2), 443-477.

Laibson, D. (2015). Why don't present-biased agents make commitments? *American Economic Review Paper and Proceedings*, *105*(5), 267-272.

Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet, 380* (9838), 219-229.

Loewenstein, G., Brennan, T., & Volpp, K. G. (2007). Asymmetric paternalism to improve health behaviors. *JAMA, 298*(20), 2415-2417.

Mainous, A. G., III, Tanner, R. J., Anton, S. D., Jo, A., & Luetke, M. C. (2017). Physical activity and abnormal blood glucose among healthy weight adults. *American Journal of Preventive Medicine 53*(1), 42-47.

Mantzari, E., Vogt, F., Shemilt, I., Wei, Y., Higgins, J. P., & Marteau, T. M. (2015). Personal financial incentives for changing habitual health-related behaviors: A systematic review and meta-analysis. *Preventive Medicine, 75*, 75-85.

Marteau, T. M., Ashcroft, R. E., & Oliver, A. . (2009). Using financial incentives to achieve healthy behaviour. *BMJ: British Medical Journal (Online), 338*(b1415).

Nicklas, J. M., Huskey, K. W., Davis, R. B., & Wee, C. C. (2012). Successful weight loss among obese U.S. adults. *American Journal of Preventive Medicine, 42*(5), 481-485.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review, 89*(1), 103-124.

Patel, M. S., Asch, D. A., Roy Rosin, M., Small, D. S., Bellamy, S. L., Heuer, J., . . . Volpp, K.G. (2016). Framing financial incentives to increase physical activity among overweight and obese adults. *Annals of Internal Medicine, 164*(6), 385-394.

Rhodes, R. E., & Dickau, L. (2012). Experimental evidence for the intention–behavior relationship in the physical activity domain: A meta-analysis. *Health Psychology. 31*(6), 724-727.

Richard, R., De Vries, N. K., & Van der Pligt, J. (1998). Anticipated regret and precautionary sexual behavior1. *Journal of Applied Social Psychology, 28*(15), 1411-1428.

Rogers, T., Milkman, K. L., & Volpp, K. G. (2014). Commitment devices: Using initiatives to change behavior. *JAMA, 311*(20), 2065-2066.

Rohde, K. I. M., & Verbeke, W. (2017). We like to see you in the gym—A field experiment on financial incentives for short and long term gym attendance. *Journal of Economic Behavior & Organization, 134*, 388-407.

Royer, H., Stehr, M. F., & Sydnor, J. R. (2012). Incentives, commitments and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company. *American Economic Journal - Applied Economics 7*(3), 51-84.

Schlam, T. R., Wilson, N. L., Shoda, Y., Mischel, W., & Ayduk, O. (2013). Preschoolers' delay of gratification predicts their body mass 30 years later. *The Journal of Pediatrics, 162*(1), 90-93.

Soman, D., Ainslie, G., Frederick, S., Li, X., Lynch, J., Moreau, P., . . . Trope, Y. (2005). The psychology of intertemporal discounting: Why are distant events valued differently from proximal ones? *Marketing Letters, 16*(3-4), 347-360.

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies, 23*(3), 165-180.

Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of political Economy*, *112*(S1), S164-S187.

Thompson, W. R. (2015). Worldwide survey of fitness trends for 2016: 10th Anniversary Edition. *ACSM's Health & Fitness Journal*, *19*(6), 9-18.

Tiessen-Raaporst, A. (2015). *Rapportage sport 2014*. Den Haag: Sociaal en Cultureel Planbureau.

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Schipper, M., Zeelenberg, M., Polder, J. J., & Prast, H. M. (2016). Design and protocol of the weight loss lottery- a cluster randomized trial. *Contemporary Clinical Trials*, *49*, 109-115.

Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial incentive–based approaches for weight loss: A randomized trial. *JAMA*, *300*(22), 2631-2637.

Wally, C. M., & Cameron, L. D. (2017). A randomized-controlled trial of social norm interventions to increase physical activity. *Annals of Behavioral Medicine*, 1-10.

Weller, R. E., Cook, E. W., Avsar, K. B., & Cox, J. E. (2008). Obese women show greater delay discounting than healthy-weight women. *Appetite*, *51*(3), 563-569.

Wertenbroch, K. (1998). Consumption self-control by rationing purchase quantities of virtue and vice. *Marketing science*, *17*(4), 317-337.

WHO, World Health Organization (2010). *Global recommendations on physical activity for health*. Geneva: WHO.

WHO, World health Organization. (2011). *Action plan for implementation of the European Strategy for the Prevention and Control of Noncommunicable Diseases 2012–2016*. Copenhagen: WHO Regional Office for Europe.

WHO, World Health Organization. (2015). *Global Health Observatory data*

repository. Retrieved from http://apps.who.int/gho/data.

Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making*, *12*(2), 93-106.

Zeelenberg, M., & Pieters, R. (2004). Consequences of regret aversion in real life: The case of the Dutch postcode lottery. *Organizational Behavior and Human Decision Processes*, *93*(2), 155-168.

Zeelenberg, M., & Pieters, R. (2007). A theory of regret regulation 1.0. *Journal of Consumer Psychology*, *17*(1), 3-18.

CHAPTER 5

Physical activity after commitment lotteries: examining long-term results in a cluster randomized trial

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Schipper, M., Zeelenberg, M., Berkhout, S., Polder, J.J., & Prast, H. M. (2018). Physical activity after commitment lotteries: examining long-term results in a cluster randomized trial. *Journal of Behavioral Medicine* 41(4), 483-493.

ABSTRACT

To overcome self-control difficulties, people can commit to their health goals by voluntarily accepting deadlines with consequences. In a commitment lottery, the winners are drawn from all participants, but can only claim their prize if they also attained their gym-attendance goals. In a 52-week, three-arm trial across six company gyms, we tested if commitment lotteries with behavioral economic underpinnings would promote physical activity among overweight adults. In previous work, we presented an effective 26-week intervention. In the present paper we analyzed maintenance of goal attainment at 52-week follow-up and the development of weight over time. We compared weight and goal attainment (gym attendance ≥ 2 per week) between three arms that –in the intervention period- consisted of (I) weekly short-term lotteries for 13 weeks; (II) the same short-term lotteries in combination with an additional long-term lottery after 26 weeks; and (III) a control arm without lottery-deadlines. After a successful 26-week intervention, goal attainment declined between weeks 27 and 52 in the long-term lottery arm, but remained higher than in the control group. Goal attainment did not differ between the short-term lottery arm and control arm. Weight declined slightly in all arms in the first 13 weeks of the trial and remained stable from there on. Commitment lotteries can support regular gym attendance up to 52 weeks and more research is needed to achieve higher levels of maintenance and weight loss.

The World Health Organization has identified physical inactivity as the fourth leading risk factor for global mortality, accounting for 3.2 million deaths annually (Forouzanfar et al., 2016). While many people know that regular physical activity (PA) is beneficial for their health and can contribute to weight management (Hildebrandt et al., 2007), most Americans (79%) and Europeans (66%) do not meet recommended levels of PA (CDC, 2014; Lee et al., 2012; EC, 2014). Besides, overweight and obese individuals generally exercise less than normal-weight individuals (CBS & RIVM, 2017) and the prevalence of obesity has more than doubled since 1980 (WHO, 2015): approximately 70% of Americans and 62% of Europeans are currently overweight or obese. Consequently, research in the field of health promotion has yielded numerous effective ways for people to change their health behaviors (Hoeymans et al., 2014). Still, most short-term interventions only have short-term effects (Van den Berg & Schoemaker, 2010). Although many people intend to improve their health by exercising on a regular basis for longer periods, the majority fails to follow through (Kooreman & Prast, 2010).

The progressing field of behavioral economics has identified systematic and predictable decision patterns that can explain why people make decisions that deviate from their own long-term health goals. Failures of self-control have been associated with present bias: the human tendency to disproportionally overweigh costs and benefits that are immediate (e.g., exercising vs. relaxing) over those that are delayed (e.g., good health in the future; Ainslie, 1975; Laibson, 1997). Consequently, people often intend to exercise on a regular basis, but eventually attend their gym less frequently than they had planned to (DellaVigna & Malmendier, 2006; Schumacher et al., 2017). In the present work, we aimed to test whether some of the same decision-biases that contribute to unhealthy behaviors can be used to durably assist individuals who have trouble sticking to their PA goals.

There is increasing behavioral economic evidence for the notion that people who foresee their self-control troubles can benefit from -and are willing to use- interventions that are known as commitment devices (Bryan et al., 2010). Commitment devices are defined as voluntarily imposed arrangements that restrict future behavior to avoid temptation (Rogers et al., 2014). For example, people cut up their credit cards or embrace withdrawal penalties on their savings account to avoid overspending and undersaving in the future (Beshears et al., 2015; Bryan et al., 2010) or request nearby deadlines for their work to preempt procrastination (Ariely & Wertenbroch, 2002). Likewise, Dutch employees are happy with mandatory pension savings because they fear that otherwise they would not save enough for retirement (Van Rooij et al., 2007).

Psychologically grounded commitment devices have also been tested as a tool to support health behavior change (Rogers et al., 2014). Based on the principle that humans dislike losses more than they like gains of equal size (Kahneman & Tversky, 1979), individuals have agreed to forfeit their monetary deposits at voluntarily imposed deadlines if they do not quit smoking

or fail to stick to their diet (Giné et al., 2010; Halpern et al., 2012; Volpp et al., 2008). Recently, lottery-based commitment devices have also been demonstrated to support weight loss, medication adherence and walking (Kimmel et al., 2012; Patel et al., 2016; Volpp et al., 2008). The lotteries aimed to tap into the human tendency to avoid regret (Christy et al., 2016; Ferrer et al., 2012; Zeelenberg & Pieters, 2007) by only awarding prizes to lottery winners who attained their health goals and informing unsuccessful lottery winners on their forgone prizes.

For the present trial, we utilized previous psychological and behavioral economic knowledge to design lottery deadlines aimed at assisting overweight adults in attaining their weekly gym attendance goals. In our *commitment lotteries* we also used the guarantee of feedback by only awarding prizes to lottery winners who attained their attendance goals and informing unsuccessful lottery winners on their forgone prizes as a way to emphasize the deadlines.

As we presented in previous work (Van der Swaluw et al., 2018), in the current sample, 13 weekly lotteries (short-term lottery arm) supported regular gym attendance. After the 13 weekly lotteries, participants in our trial attended their gym considerably less. Adding an additional long-term lottery deadline after 26 weeks (long-term lottery arm) partly averted the decline in gym attendance, indicating that a long-term lottery deadline can help sustain regular gym attendance up to 13 additional weeks. The adjusted probability of goal attainment (week-gym attendance $\geq$ 2) in weeks 1 to 13 was 57% in the short-term lottery arm, 66% in the long-term lottery arm and 25% in the control arm. In both lottery arms, 8 out of the 13 weekly winners were eligible for their prize. Between weeks 14 to 26, the adjusted probability of week-goal attainment was highest in the long-term lottery arm (51%) versus 23% in the other arms. The majority of participants (55%) in the long-term lottery arm was eligible to claim the prize in week 26 if they would have won it.

For the present paper, we analyzed gym attendance after all lotteries ended. Although there is evidence that commitment devices (e.g., voluntary imposed deadlines with consequences) in the health domain are effective in the short-run, the long-term effects are often unsatisfactory (John et al., 2011; Patel et al., 2016; Patel et al., 2016b; Royer et al., 2012). Consequently, one of the key challenges in the application of commitment devices is either safeguarding maintenance of behavior change after an intervention or continuing its application (Halpern et al., 2012; Rogers et al., 2014). We focused on the former and with this purpose we examined individual goal attainment (week-gym attendance $\geq$ 2) after the interventions.

After all deadlines had passed, gym attendance was monitored up to week 52 in all arms to examine maintenance of behavior change. Specifically, we questioned whether goal attainment differed between the long-term lottery arm, the short-term lottery arm and the control arm between weeks 27 and 52. We hypothesized that goal attainment in the lottery arms would

decline between weeks 27 and 52, but still be significantly higher than goal attainment in the control arm and would be highest in the long-term lottery arm (Van der Swaluw et al., 2016).

We additionally studied weight patterns over the course of the trial. Next to –among others– genetic, sociocultural, economic and environmental factors, modifiable behaviors (diet and PA) markedly contribute to the development of overweight and obesity (Bray et al., 2016). As such, regular PA can contribute to weight loss and weight management (Fogelholm, 2010; Fogelholm & Kukkonen-Harjula, 2000). Furthermore, overweight has been associated with self-control problems (Ikeda et al., 2010; Schlam et al., 2013), while commitment lotteries aim to combat this. Hence, we also explored the effect of the interventions on the development of weight over the 52 weeks of this trial.

METHOD

Interventions

We compared two intervention arms to a control arm. Participants in all arms set the goal to exercise twice per week. Participation was free of charge and all participants were offered monthly statistics on their progress. The control arm was also actively monitored, but was neither aware of- nor participated in the lotteries. In all arms, participants were supervised as usual by the gym staff and were free to choose their preferred mode of exercise. All gyms were equipped to facilitate endurance training, strength training and standardized group classes (e.g., circuit training). As a normal part of the gym membership, participants had access to a variety of ready-to-use training schedules that fit different exercise goals (e.g., weight loss vs. enhancing stamina).

Short-term lottery arm

Participants in the short-term lottery arm participated in 13 weekly lotteries worth € 100 each of the first 13 weeks of the trial. The weekly winners were randomly drawn out of all participants in this arm and were communicated to all by email and text message. The weekly winners were only eligible for their prize if they attended their gym at least twice that week (the week goal). Of key importance was also that lottery winners who did not attain their goal were informed about their forfeited prize. If a participant won one of the weekly lotteries, but did not attain the week goal in that week, it was communicated that "you won the lottery this week, but cannot claim your prize, since you did not meet your goal of attending the gym twice". Participants were fully informed and reminded about the possibility of this counterfactual feedback and the course of the lotteries. All other candidates in this arm were informed whether the prize was granted or forfeited. In each of the 13 weeks, the winner was drawn out of all participants regardless of prior performance. Thus, every new week meant a renewal of the commitment to exercise twice a week.

Long-term lottery arm

The first 13 weeks in the long-term lottery arm were identical to the short-term lottery arm. Participants knew prior to the trial that weeks 14 to 26 would also be part of the intervention. After 26 weeks a luxury family-vacation voucher was ascribed to a randomly drawn participant in this arm. Again, the winner was communicated to all participants by email and text message. Participants knew and were reminded that they would always learn the outcome of the lottery, but that the prize could only be claimed if the winner had attained his or her goal in at least 9 of the second 13 weeks (70% between weeks 14 and 26). We guaranteed that the prize would eventually be awarded: if the winner was not eligible for the prize, he or she would be informed about the forgone prize and another winner would be drawn until the prize could be awarded according to the rules mentioned above.

Design and setting

The study design and details on randomization, blinding, eligibility, recruitment and measurement protocols have been published before (Van der Swaluw et al., 2016). In brief, we set up a three-arm, parallel group, cluster randomized trial running for 52 weeks with 163 participants in six company gyms (clusters) across the Netherlands. Figure 1 displays the trial flow over 52 weeks. The gyms were branches of international fitness agency High Five, which provides corporate fitness training in 36 organizations across the Netherlands. The trial was reviewed and approved by the Tilburg University Ethical Review Board (EC-2014.42a) and is registered in the Dutch Trial Register (NTR5559). The lottery drawings were performed by the independent Game Management Department of the Dutch State Lottery under supervision of a notary.

The six gyms were randomized to one of three arms. As such, every arm contained two gyms. Participants were eligible if they expressed in a survey their goal to exercise twice or more per week, were between the ages of 18-65, were overweight ($25 \leq BMI < 40$) and had not planned a leave of absence of more than 4 weeks in the first 26 weeks of the trial. Participants were blinded from the other trial arms. Informed consent was obtained from all individual participants included in the study. Table 1 displays the baseline data of participants in all three arms.

Figure 1: Study design and flow of gyms and participants.

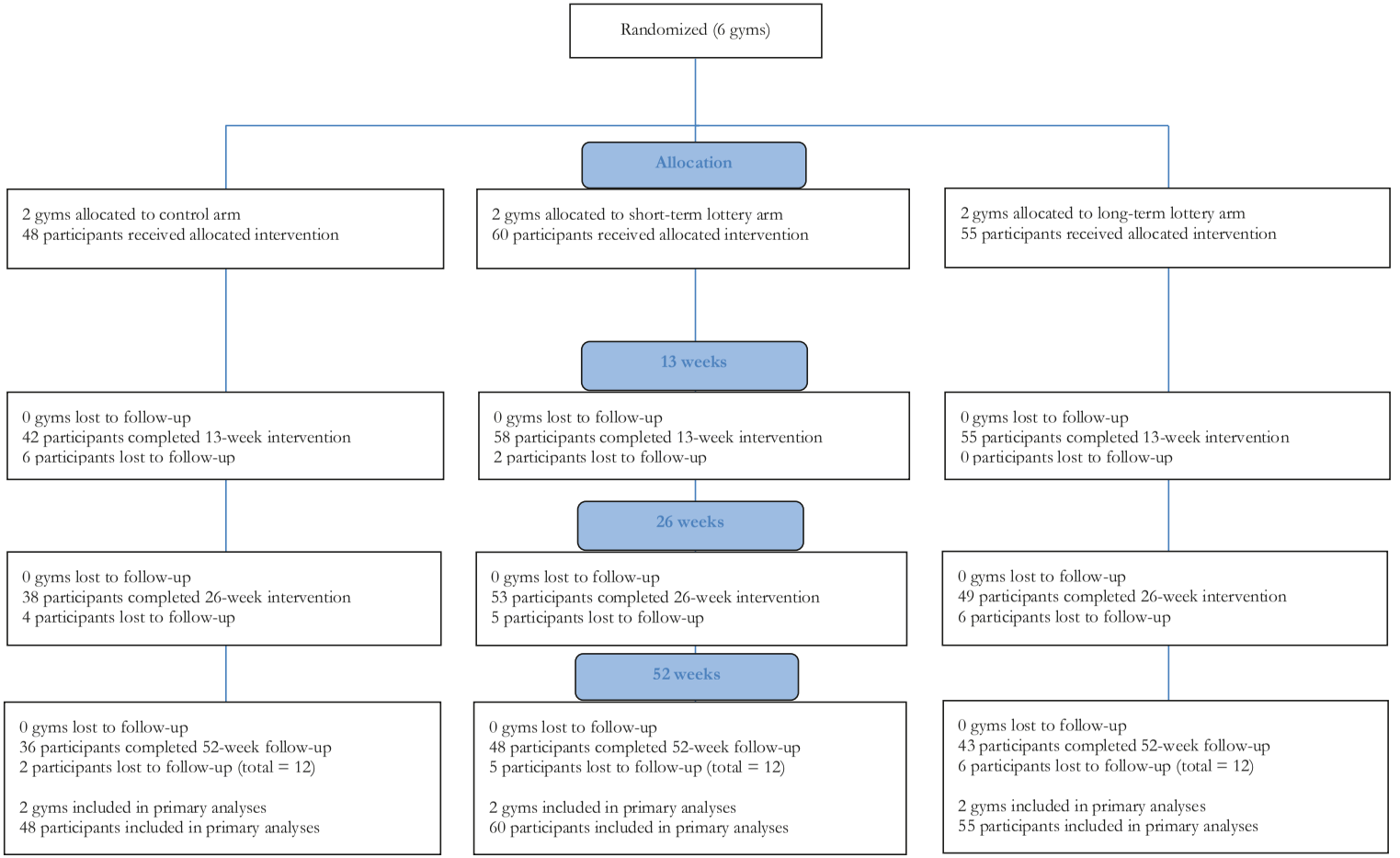


Table 1. Participant demographics per study arm

Characteristic	control (n = 48)	short-term lotteries (n = 60)	long-term lottery (n = 55)
Age, mean (SD)	50 (9.84)	49.3(9.33)	45(9.58)
Gender, no. (%)			
Female	16 (33.3)	21 (35)	13 (23.6)
Male	32 (66.7)	39 (65)	42 (76.4)
No survey response, no. (%)	3 (6.25)	0 (0)	1 (1.82)
Nationality, no. (%)			
Dutch	36 (80)	52 (86.7)	52 (94.5)
Other	12 (20)	8 (13.3)	3 (5.5)
Education, no. (%)			
Pre-vocational education	3 (7.9)	7 (11.5)	4 (7.3)
Pre-university education	3 (6.7)	2 (3.3)	10 (18.2)
Senior vocational training	11 (24.4)	20 (33.3)	5 (9.1)
Vocational colleges	19 (42.2)	15 (25)	23 (41.8)
University education	9 (20)	15 (25)	10 (18.2)
Other	0 (0)	1 (1.7)	2 (3.6)
Monthly net income, no. (%)			
<€1000	0 (0)	0 (0)	1 (1.8)
€1000 to €2000	10 (20.8)	6 (10)	3 (5.5)
€2000 to €3000	19 (39.6)	32 (53.3)	24 (43.6)
€3000 to €4000	8 (16.7)	15 (25)	19 (34.5)
€4000 to €5000	2 (4.2)	1 (1.7)	2 (3.6)
€5000 tot €6000	0 (0)	2 (3.3)	1 (1.8)
> €6000	1 (2.1)	0 (0)	0 (0)
Did not wish to answer	5 (10.4)	4 (6.7)	4 (7.3)
Baseline gym attendance*, mean (SD)	1.82 (0.88)	1.46 (1.17)	1.55 (1.04)
Weight, mean (SD)	90.14 (14.38)	96.12 (14.12)	96.6 (13.94)
BMI, mean (SD)	28.9 (3.20)	30.4 (3.73)	30.19 (3.47)
Obese, no.(%)	13 (27.1)	23 (38.3)	26 (47.3)

*Participants answered the question; "On average, how often per week did you attend the gym in the last two months?"

Outcomes & measures

Of primary interest in this trial was goal-attainment (week-gym attendance ≥ 2) measured at the participant level. We provided all gyms with an iPad connected via Wi-Fi, which allowed us to monitor attendance in real time. Throughout the 52 weeks of the trial, participants were required to check in at the iPad with their name or three digit study-ID when entering their gym. All gyms were provided with identical scales to measure weight (KERN™; 0.1% precision). Upon registering their attendance, participants were asked to weigh (without shoes) and to enter their weight (kilograms, 1 decimal) into the iPad. Participants could also select the option; "I already entered my weight this week", or "I will enter my weight later this week". Hence, weight was assessed on a weekly basis. At baseline, 13-, 26- and 52 weeks, participants were supervised by the gym personnel in entering their weight. Baseline attendance levels and demographics were assessed via an additional online questionnaire.

Analyses

Participants were the primary unit of inference in all analyses. Analyses followed the intention-to-treat principle and were conducted in R version 3.4.0 with statistical significance set at $p < 0.05$. Planned analyses can also be found in the trial protocol (Van der Swaluw et al., 2016).

To evaluate the effect of the interventions on goal attainment per week, we performed three multi-level logistic regression analyses with goal attainment from weeks 1 to 13, 14 to 26, and 27 to 52 as the dependent variable respectively. These measurements are nested within participants who are clustered within gyms. The three trial arms, time, self-reported baseline attendance, age and sex were entered as fixed effects. Random intercepts were added for both the participants and the gyms. A random slope for time on the participant level was also included in the model and hereby allowed for different time patterns among individuals. Week 42 of the trial was excluded from the analyses, because gyms were closed in that week as a result of the new-year holiday season.

When there are few gyms per arm, treatment effects may correlate with intra cluster (gym) effects. By adding the random intercept for gyms, the model estimated the treatment effects, while accounting for the clustered data pattern in gyms. Additionally, sensitivity analyses were performed by excluding each gym from the models once and comparing effects of the reduced data models to the full model.

To assess the effect of the interventions on weight in each trial period, multi-level linear regression analyses were performed with weight (kilograms) as the dependent variable. Linear multi-level modelling has shown to be a reliable technique to handle missing longitudinal outcome data (Peters et al., 2012) and was used to fit weight patterns over time, despite missing outcome measurements of participants who did not enter their weight that week. As before, these measurements are nested within participants who are clustered within gyms. Again, random intercepts were added for both the participants and the gyms. A random slope on the participant level was also added. The three trial-arms, time, baseline BMI, age and sex were included in all models as fixed effects.

RESULTS

Goal attainment

Figure 2 displays unadjusted percentages of goal attainment per arm per week. Results on goal-attainment in the first 26 weeks of the trial have been presented before (Van der Swaluw et al., 2018). Between weeks 27- 52, the aggregated percentage of goal attainment was 24% in the long-term lottery arm and 16% and 15% in the short-term arm and control arm respectively. Table 2 displays the time-adjusted odds ratios of goal attainment in each trial period.

In the logistic mixed model fitting weekly goal attainment between weeks 27-52, the long-term lottery arm had significantly higher odds of goal attainment than the control arm (OR: 7.88, 95% CI: 1.18 - 52.51, $p = .03$) and non-significant higher odds than the short-term lottery arm (OR: 4.31, 95% CI: 0.83 – 22.36, $p = .08$). The difference in goal attainment between the short-term lottery arm and the control arm was not statistically significant (OR: 1.84, 95% CI: 0.30 – 11.54, $p = .15$).

Table 2. Logistic mixed models describing goal attainment (week-attendance ≥2)

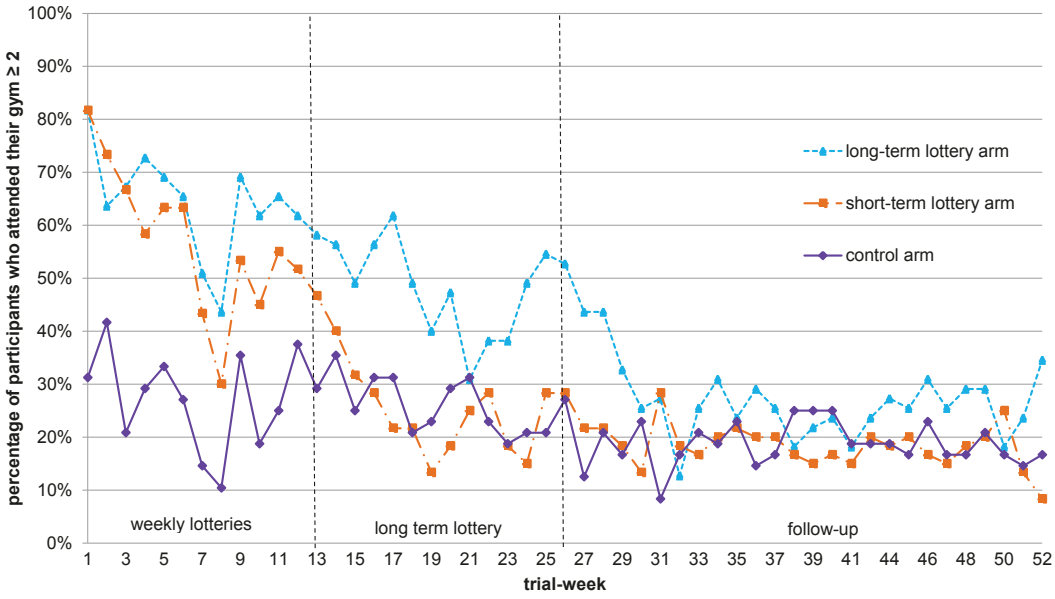
	Weeks 1-13		**Weeks 14-26**		**Weeks 27-52**	
	Odds Ratio (95% CI)		Odds Ratio (95% CI)		Odds Ratio (95% CI)	
Trial characteristics						
Control arm	(ref.)					
Short-term lotteries	12.10**	(2.54- 57.53)	1.15	(0.23 - 5.73)	1.84	(0.30 - 11.54)
Long-term lottery	13.47**	(2.76 - 65.74)	6.13*	(1.18 - 31.92)	7.88*	(1.18 - 52.51)
Time (week)	0.93	(0.86 - 1.00)	0.92	(0.84 - 1.02)	0.96	(0.91 - 1.01)
Short-term ĺ time	0.92	(0.84 - 1.01)	1.01	(0.90 - 1.13)	0.99	(0.94 - 1.04)
Long-term ĺ time	1.01	(0.91 - 1.11)	1.01	(0.91 - 1.11)	0.97	(0.92 - 1.03)
Participant characteristics						
Baseline attendance	1.39*	(1.05 - 1.85)	1.92***	(1.32 - 2.80)	2.35*	(1.40 - 3.93)
Age	1.00	(0.97 - 1.03)	1.04	(1.00 - 1.09)	1.05	(0.99 - 1.11)
Male (ref.) vs. Female	0.39**	(0.21 - 0.75)	0.61	(0.26 - 1.43)	0.76	(0.22 - 2.58)

* significant at $p < .05$
** significant at $p < .01$
*** significant at $p < .001$
Model accounts for clustered measures within gyms, participants and for temporal trends by week. Outcome is a binary term (0 or 1). CI = confidence interval. Ref = reference category.

Sensitivity analyses

The intervention effect between weeks 27 and 52 was only sensitive to the sequential exclusion of control gyms from the models. If we excluded the best performing gym from the control arm, the intervention effect of the short-term lottery arm became statistically significant. If we excluded the least performing gym from the control arm, the effect of the long-term lottery arm remained qualitatively similar, but was no longer statistically significant.

Figure 2. Goal attainment (week-gym attendance ≥ 2) over time per arm



Weight

Table 3 displays the output of the linear mixed models describing weight in each trial period. In the first 13 weeks of the trial, weight declined slightly over time in the control arm (B: -0.09, SE: 0.03, $p = .002$), the short-term lottery arm (B: -0.08, SE: 0.02, $p = .001$) and the long-term lottery arm (B: -0.06, SE: 0.03, $p = .02$). The decline in weight did not differ between arms. Between weeks 14-26 and weeks 27-52, the models display neither a significant decline of weight over time, nor significant differences between arms.

Table 3. Linear mixed models describing weight (kilograms).

	Weeks 1-13		Weeks 14-26		Weeks 27-52	
	Unstandardized Beta (SE)	*t*-value	Unstandardized Beta (SE)	*t*-value	Unstandardized Beta (SE)	*t*-value
Trial characteristics						
Control arm	-1.02 (5.47)	-0.19	-0.19 (6.26)	-0.03	-1.33 (6.79)	-0.20
Short-term lotteries	0.90 (1.85)	0.49	-0.05 (1.86)	-0.03	-0.63 (2.08)	-0.30
Long-term lottery	1.40 (1.92)	0.73	1.11 (1.94)	0.57	0.44 (2.13)	0.21
Time (week)	-0.09 (0.03)**	-3.03	-0.05 (0.03)	-1.41	-0.01 (0.04)	-0.28
Short-term X time	0.01 (0.04)	0.31	0.04 (0.05)	0.93	0.01 (0.05)	0.20
Long-term X time	0.03 (0.04)	0.80	0.05 (0.04)	1.14	-0.01 (0.05)	-0.14
Participant characteristics						
Baseline BMI	2.84 (0.18)***	15.87	2.79 (0.21)***	13.60	2.89 (0.22)***	12.94
Age	0.07 (0.07)	1.05	0.09 (0.08)	1.22	0.11 (0.09)	1.21
Male (ref.) vs. Female	13.73 (1.39)***	9.89	14.06 (1.54)***	9.16	13.13 (1.74)***	7.57

* significant at $p < .05$
** significant at $p < .01$
*** significant at $p < .001$
Model accounts for clustered measures within gyms, participants and for temporal trends by week. SE= Standard Error. Ref= reference category.

DISCUSSION

The 52-week follow-up analyses of our cluster randomized trial show moderately sustained levels of goal attainment (gym week-attendance ≥ 2) six months after completing a 26-week lottery intervention. Up to one year after the start of the intervention, participants who entered 13 weekly commitment lotteries (weeks 1-13), followed by an additional lottery 13 weeks later (weeks 14-26) were more likely to attend their gym twice per week than participants in the control arm. In the same follow-up period, goal attainment in the short-term lottery arm (13 weekly commitment lotteries) did not differ significantly from the control arm or long-term lottery arm. Weight declined slightly in the first 13 weeks of the trial in all arms and remained stable from there on.

The present study may contribute to the pursuit of methods for sustainable behavior change. Commitment lotteries have been used relatively sporadically in the field of health promotion. Starting a decade ago, Volpp et al. (2008) have offered lotteries to overweight participants to support their weight loss attempts. While their lotteries were effective for 16 weeks, participants regained weight after the intervention. Likewise, Patel et al. (2016) have effectively used team-based lotteries to stimulate walking for 13 weeks, but differences between intervention and control deteriorated during the 26-week follow-up period. Although there is evidence for their short-term effectiveness, the longevity of commitment lotteries is unsatisfactory. In the present trial, we observe a significantly higher likelihood of goal attainment in the long-term lottery

arm than control in the six months after the intervention. The observation that the long-term lottery-arm does not underperform the control arm during follow-up implies that the 26-week intervention has had a net effect on gym attendance over a 1-year period.

Still, in the interpretation of our results, some nuance is warranted. In Figure 2 it can be observed that levels of goal attainment in the long-term lottery arm dropped after the 26-week lottery deadline. Accordingly, average week-goal attainment in the long-term lottery arm halved during follow-up, relative to weeks 14 to 26. Therefore, it cannot be concluded that participants in the long-term lottery arm remained equivalently committed to their gym attendance goal as before. Similar to previous studies, behavior change is shown to be challenging to maintain after commitment lotteries end.

Nonetheless, the tested combination of short-term lotteries and a long-term lottery in this trial provides novel and useful insights that can be built upon in future trials to help optimize the long-term effectiveness of commitment lotteries. For example, for the winner to be eligible for the long-term lottery prize, he or she had to attain at least 9 week-goals between weeks 14 to 26. The majority of participants in the long-term lottery arm did so. Therefore, future studies can experiment by lengthening the long-term lottery deadline (e.g., to 36 weeks; John et al., 2011) to test if this will further stimulate sustained goal attainment. This exploration may result in a deadline that is nearby enough to be salient in the present, while being lengthy enough to promote sustained behavior change after the deadline.

Similar to previous research, behavioral economic and psychological insights on decision-making could be well incorporated in health promotion (Loewenstein et al., 2012). First, to leverage present bias (the overweighing of the present), we imposed nearby deadlines to draw the consequences of procrastination nearer (Ariely & Wertenbroch, 2002). Second, because people tend to overestimate small probabilities, a lottery is an effective and scalable tool to make missing the deadlines potentially costly for participants (Kahneman & Tversky, 1979). Third, we leveraged the human tendency to avoid regret by drawing the lottery prize from all participants and informing non-eligible winners on their forgone prize (Zeelenberg & Pieters, 2004, 2007). Fourth, while we aimed to stimulate long-term behavior change, we used week goals to fit individuals' impatience and to facilitate the human tendency of using of temporal landmarks (e.g., Mondays) to relegate misfortune to the past and have a fresh start (Dai et al., 2014; O'Donoghue & Rabin, 1999).

The results of this trial also offer several practical insights for health professionals, policy makers, insurers or employers who aim to support people in achieving their health goals. Next to its effectiveness, the costs of prevention are considered a key aspect in the determination of its value to businesses and society (Van den Berg & Schoemaker, 2010). The use of psychological and behavioral economic insights in the design of the lotteries can enhance the

psychological impact of money that is spent and could offer the opportunity to implement commitment lotteries at low costs. To stimulate regular gym attendance up to 52 weeks in the long-term lottery arm, we spent only €2.21 per participant per week (awarded prizes ÷ participants ÷ weeks). In perspective; this is 0.6% of the Dutch minimum week wage (€361.25). The costs in future applications may be further reduced if participants are also willing to pay for their wish to commit, similar to self-imposed withdrawal penalties on savings accounts (Beshears et al., 2015) or betting one's money on personal health goals with a deposit contract (Halpern et al., 2012). It would be valuable to explore the characteristics of participants and potential organizers of commitment lotteries that contribute to the optimal balance between attractiveness and effectiveness.

Despite their potential, commitment devices remain underused (Halpern et al., 2012; Rogers et al., 2014) - a fact that might be explained by the range of open issues on their implementation. For example, employees or patients outside the interventions may resent others receiving lottery prizes (Loewenstein et al., 2012) (which may also be circumvented by asking participants to pay for lottery tickets). Besides, commitment lotteries were especially effective while they were active and -analogous to drugs- lost most of their effectiveness once people stopped 'taking them'. This finding stimulates thinking about the type of commitment device that maintains its impact over repeated (and possibly infinite) application. While goal attainment declined over time in our trial (see Figure 2), the decline of goal attainment over time did not differ significantly between the lottery arms and the control arm. Therefore, it would be interesting to design and test similar commitment lotteries that, like many commercial lotteries, are endlessly repeated and accessible.

Another issue surrounding the applicability of commitment lotteries is their target population. The majority of our sample was male (69.3%), which was similar to the general population of the six gyms (69.8% male). Nonetheless, future studies would benefit from enrolling an even proportion of females and males. A meta-analysis by Haff et al. (2015) that compared the effectiveness of several commitment lotteries between demographic groups found no gender differences, nor differences in education or broad ranges of income (except a slight reduced effect with incomes > $87.500). In previous analyses, we also observed no differences in intervention effectiveness between income- or education categories (Van der Swaluw et al., 2018). This may be an indication of commitment lotteries being broadly applicable. A next step in the development of commitment devices could be the assessment of design features that contribute to their efficiency and feasibility across different populations. Hence, although there is increasing evidence for the effectiveness of commitment lotteries, more applied research could enhance their generalized practicability.

Commitment lotteries aimed at promoting regular PA did not result in substantial weight loss. After a moderate decline in weeks 1 to 13, weight remained stable over time in all arms, which

is contrary to projections of overweight and obesity progressing over time (Wang et al., 2008). Considering that medical complications and associated health care costs rise progressively as BMI increases (Cawley & Meyerhoefer, 2012), stable weight is not all bad. Nonetheless, no meaningful weight loss was achieved as a result of increased PA, meaning that most participants remained exposed to the increased risks of cardiovascular diseases, diabetes type 2 and cancers that are associated with overweight and obesity (Forouzanfar et al., 2016; WHO, 2013). Our findings are in line with the general conception that regular exercise can contribute to weight loss but often not solely (Fogelholm, 2010; Fogelholm & Kukkonen-Harjula, 2000). In some occasions, increases in PA have also been found to result in weight gain by calorie compensation (McCaig et al., 2016).

To achieve significant and sustained weight loss, regular PA is only one part of the multicomponent (lifestyle training, nutrition and PA) interventions that are acknowledged to be effective (Bray et al., 2016). Still, weight loss interventions that include an exercise component are often more effective than interventions that do not (De Roon et al., 2017; Jeffery et al., 2003). Given that overweight individuals generally exercise less than normal weight individuals (CBS & RIVM, 2017), any multicomponent weight loss intervention that also encompasses PA might still benefit from including commitment lotteries for PA.

This trial is subject to several limitations. First, while it is clear that commitment lotteries did not promote meaningful weight loss, it remains unclear why this was the case in this trial (e.g., due to the type of exercise, calorie compensation, or loss of fat that may have been compensated by increased muscle mass). We chose to focus on gym attendance as our main dependent variable, rather than weight loss, because this is under more direct volitional control by the participants. In future instances, it may be valuable to explore the need for additional weight loss guidance when participants enter commitment lotteries for PA.

A second limitation is that our trial limited the promotion of regular PA to gym attendance, while activities with lower intensity (e.g., walking and recreational cycling) are also known to promote health outcomes (Fogelholm, 2010). The benefit of our approach is the novel and scalable context in which commitment lotteries were shown to be effective (company gyms), while exercise was supervised by the gym personnel.

Third, our trial included no more than six gyms and randomization at the gym level increased the potential influence of intra gym effects in the effectiveness of the interventions. The sensitivity analysis (where we excluded the least performing control-gym from the model) reduced the sustained effect of the long-term lottery to non-significance. We accounted for intra-gym influences in our multi-level analyses and randomization at the gym level had the benefit that treatment contamination was minimized. Still, future studies can potentially avoid these issues by enrolling more gyms.

Conclusion

Regular PA has numerous health benefits. While many people aim to exercise on a regular basis, there are multiple predictable behavioral patterns that hamper the progression of a health-intention to sustained behavior. Effective and scalable support of long-term regular exercise is needed. Commitment lotteries that were designed to leverage psychological knowledge on decision-making can help cope with the challenges of health behavior change at low costs. While the 26-week intervention supported regular PA up to 52 weeks, levels of gym attendance declined after all deadlines had passed. Participants did not remain equivalently committed as before the long-term lottery deadline. Future research in broader populations could reveal to what degree commitment lotteries remain effective over longer time periods or if they are endlessly accessible.

REFERENCES

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin, 82*(4), 463.

Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: self-control by precommitment. *Psychological Science, 13*(3), 219-224.

Beshears, J., Choi, J. J., Harris, C., Laibson, D., Madrian, B. C., & Sakong, J. (2015). Self control and commitment: can decreasing the liquidity of a savings account increase deposits? *National Bureau of Economic Research, w21474*, 1–53.

Bray, G. A., Fruhbeck, G., Ryan, D. H., & Wilding, J. P. (2016). Management of obesity. *The Lancet, 387*(10031), 1947-1956.

Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment devices. *Annual Review of Economics., 2*(1), 671-698.

Cawley, J., & Meyerhoefer, C. (2012). The medical care costs of obesity: an instrumental variables approach. *Journal of Health Economics, 31*(1), 219-230.

CBS, Central Bureau voor de Statistiek., & RIVM, Rijksinstituut voor Volksgezondheid en Milieu. (2017). *Gezondheidsenquête/Leefstijlmonitor.* Retrieved from statline.cbs.nl.

CDC, Centers for Disease Control and Prevention, Division of Nutrition, P. A., and Obesity. (2014). Facts about Physical Activity. Retrieved from www.cdc.gov/physicalactivity/data/facts.htm

Christy, S. M., Winger, J. G., Raffanello, E. W., Halpern, L. F., Danoff-Burg, S., & Mosher, C. E. (2016). The role of anticipated regret and health beliefs in HPV vaccination intentions among young adults. *Journal of Behavioral Medicine, 39*(3), 429-440.

Dai, H., Milkman, K. L., & Riis, J. (2014). The fresh start effect: Temporal landmarks motivate aspirational behavior. *Management Science, 60*(10), 2563-2582.

De Roon, M., Van Gemert, W. A., Peeters, P. H., Schuit, A. J., & Monninkhof, E. M. (2017). Long-term effects of a weight loss intervention with or without exercise component in postmenopausal women: A randomized trial. *Preventive Medicine Reports, 5*, 118-123.

DellaVigna, S., & Malmendier, U. (2006). Paying not to go to the gym. *The American Economic Review*, 694-719.

European Commission: Strategy, Corporate Communication Actions and Eurobarometer Unit

(2014). Sport and physical activity. *Special Eurobarometer 412.*

Ferrer, R. A., Klein, W. M. P., Zajac, L. E., Land, S. R., & Ling, B. S. (2012). An affective booster moderates the effect of gain- and loss-framed messages on behavioral intentions for colorectal cancer screening. *Journal of Behavioral Medicine, 35*(4), 452-461.

Fogelholm, M. (2010). Physical activity, fitness and fatness: relations to mortality, morbidity and disease risk factors. A systematic review. *Obesity Reviews, 11*(3), 202-221.

Fogelholm, M., & Kukkonen-Harjula, K. (2000). Does physical activity prevent weight gain – a systematic review. *Obesity Reviews, 1*(2), 95-111.

Forouzanfar, M. H., Afshin, A., Alexander, L. T., Anderson, H. R., Bhutta, Z. A., Biryukov, S., . . . Charlson, F. J. (2016). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388*(10053), 1659-1724.

Giné, X., Karlan, D., & Zinman, J. (2010). Put your money where your butt is: a commitment contract for smoking cessation. *American Economic Journal: Applied Economics, 2* (4), 213-235.

Haff, N., Patel, M. S., Lim, R., Zhu, J., Troxel, A. B., Asch, D. A., & Volpp, K. G. (2015). The role of behavioral economic incentive design and demographic characteristics in financial incentive-based approaches to changing health behaviors: a meta-analysis. *American Journal of Health Promotion, 29*(5), 314-323.

Halpern, S. D., Asch, D. A., & Volpp, K. G. (2012). Commitment contracts as a way to health. *BMJ: British Medical Journal, 344*(1), e522-e522.

Hildebrandt, V.H., Ooijendijk, W.T.M., Hopman-Rock, M.(2007). *Trendrapport bewegen en gezondheid 2004/2005*. Hoofddorp/Leiden: TNO

Hoeymans, N., Van Loon, A., Van den Berg, M., Harbers, M., Hilderink, H., Van Oers, J., & Schoemaker, C. (2014). *Een gezonder Nederland: Volksgezondheid Toekomst Verkenning 2014*. Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu.

Ikeda, S., Kang, M.-I., & Ohtake, F. (2010). Hyperbolic discounting, the sign effect, and the body mass index. *Journal of Health Economics, 29*(2), 268-284.

Jeffery, R. W., Wing, R. R., Sherwood, N. E., & Tate, D. F. (2003). Physical activity and weight loss: does prescribing higher physical activity goals improve outcome? *The American Journal of Clinical Nutrition, 78*(4), 684-689.

John, L., Loewenstein, G., Troxel, A., Norton, L., Fassbender, J., & Volpp, K. (2011). Financial incentives for extended weight loss: a randomized, controlled trial. *Journal of General Internal Medicine, 26*(6), 621-626.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society, 47*(2), 263-291.

Kimmel, S. E., Troxel, A. B., Loewenstein, G., Brensinger, C. M., Jaskowiak, J., Doshi, J. A., . . . Volpp, K. (2012). Randomized trial of lottery-based incentives to improve warfarin adherence. *American Heart Journal, 164*(2), 268-274.

Kooreman, P., & Prast, H. (2010). What Does Behavioral Economics Mean for Policy? Challenges to Savings and Health Policies in the Netherlands. *De Economist, 158*(2), 101-122.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics, 112*(2), 443-477.

Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet, 380* (9838), 219-229.

Loewenstein, G., Asch, D. A., Friedman, J. Y., Melichar, L. A., & Volpp, K. G. (2012). Can behavioural economics make us healthier? *BMJ: British Medical Journal (Online), 344, e3482*.

Van den Berg, M. & Schoemaker, C.G. (2010). *Effecten van preventie. Deelrapport van de VTV 2010 Van gezond naar beter.* Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu.

McCaig, D. C., Hawkins, L. A., & Rogers, P. J. (2016). Licence to eat: Information on energy expended during exercise affects subsequent energy intake. *Appetite*, *107*, 323-329.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, *89*(1), 103-124.

Patel, M. S., Asch, D. A., Rosin, R., Small, D. S., Bellamy, S. L., Eberbach, K., . . . Wesby, L. (2016). Individual Versus Team-Based Financial Incentives to Increase Physical Activity: A Randomized, Controlled Trial. *Journal of General Internal Medicine*, *31*, 1-9.

Patel, M. S., Asch, D. A., Roy Rosin, M., Small, D. S., Bellamy, S. L., Heuer, J., . . . Volpp, K.G. (2016)b. Framing Financial Incentives To Increase Physical Activity Among Overweight And Obese Adults. *Annals of Internal Medicine*, *164*(6), 385-394.

Peters, S. A. E., Bots, M. L., den Ruijter, H. M., Palmer, M. K., Grobbee, D. E., Crouse, J. R., . . . Koffijberg, H. (2012). Multiple imputation of missing repeated outcome measurements did not add to linear mixed-effects models. *Journal of Clinical Epidemiology*, *65*(6), 686-695.

Rogers, T., Milkman, K. L., & Volpp, K. G. (2014). Commitment devices: Using initiatives to change behavior. *JAMA*, *311*(20), 2065-2066.

Royer, H., Stehr, M. F., & Sydnor, J. R. (2012). Incentives, commitments and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company. *American Economic Journal: Applied Economics*, *7*(3), 51-84.

Schlam, T. R., Wilson, N. L., Shoda, Y., Mischel, W., & Ayduk, O. (2013). Preschoolers' Delay of Gratification Predicts their Body Mass 30 Years Later. *The Journal of Pediatrics*, *162*(1), 90-93.

Schumacher, L. M., Arigo, D., & Thomas, C. (2017). Understanding physical activity lapses among women: responses to lapses and the potential buffering effect of social support. *Journal of Behavioral Medicine*, *40*(5), 740-749.

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Schipper, M., Zeelenberg, M., Berkhout, S., Polder J.J., Prast, H. M. (2018). Commitment lotteries promote physical activity among overweight adults: a cluster randomized trial. *Annals of Behavioral Medicine* 52(4), 342-351.

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Schipper, M., Zeelenberg, M., Polder, J. J., & Prast, H. M. (2016). Design and protocol of the weight loss lottery- a cluster randomized trial. *Contemporary Clinical Trials*, *49*, 109-115.

Van Rooij, M. C., Kool, C. J., & Prast, H. M. (2007). Risk-return preferences in the pension domain: are people able to choose? *Journal of Public Economics*, *91*(3), 701-722.

Volpp, K. G., John, L. K., Troxel, A. B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial incentive–based approaches for weight loss: A randomized trial. *JAMA*, *300*(22), 2631-2637.

Wang, Y., Beydoun, M. A., Liang, L., Caballero, B., & Kumanyika, S. K. (2008). Will all Americans become overweight or obese? Estimating the progression and cost of the US obesity epidemic. *Obesity*, *16*(10), 2323-2330.

WHO, World Health Organization (2013). *Global action plan for the prevention and control of noncommunicable diseases 2013-2020*. Geneva:WHO.

WHO, World Health Organization (2015). *Global Health Observatory data repository*. Retrieved from http://apps.who.int/gho/data/

Wolfson, S., & Briggs, P. (2002). Locked into gambling: Anticipatory regret as a motivator for playing the National Lottery. *Journal of Gambling Studies*, *18*(1), 1-17.

Zeelenberg, M., & Pieters, R. (2004). Consequences of regret aversion in real life: The case of the Dutch postcode lottery. *Organizational Behavior and Human Decision Processes*, *93*(2), 155-168.

Zeelenberg, M., & Pieters, R. (2007). A theory of regret regulation 1.0. *Journal of Consumer Psychology*, *17*(1), 3-18.

CHAPTER 6

Consequences of regret aversion in intertemporal choice

Van der Swaluw, K., Lambooij, M. S., Zeelenberg, M., Mathijssen, J. J. P., Polder, J.J., & Prast, H. M. (2018). Consequences of regret aversion in intertemporal choice. Under review

ABSTRACT

If decision-makers expect to learn the outcome of a non-chosen decision-alternative (i.e., expect to receive counterfactual feedback), this impacts their decisions. It remained unclear to what extent such feedback retains its influence if feedback is delayed and whether feedback influences delay discounting. In three experiments (N = 176; N = 592; N = 340 respectively), we varied counterfactual feedback on delayed and uncertain gains to test if this would affect I) certainty equivalents and II) degrees of delay discounting. Counterfactual feedback did not attenuate delay discounting, but did affect decisions in the present with delayed consequences. In a choice between an immediate certain gain and a delayed gamble, participants who would always learn the outcome of the delayed gamble valued the gamble higher than participants without feedback. This pattern was not significant when amounts were higher. Counterfactual feedback may be effective in supporting present decisions with delayed uncertain consequences.

'Reality is a cloud of possibility'
–Daniel Kahneman (in Lewis, 2016. p.312).

In decision-making under uncertainty, the desirability of each outcome is related to the expected pain and pleasure associated with that outcome. Logically, outcomes with more expected joy and less expected misery are more desirable (Von Neumann & Morgenstern, 1944; 2007). The desirability of an outcome is further determined by how it compares to the outcome of the non-chosen, rejected alternative. If the non-chosen alternative would have resulted in a better outcome, people feel regret, and this negatively affects the evaluation of the obtained outcome. People may take this into account when deciding and choose in such a way to avoid the regret that they feel when the chosen alternative ends up being worse than the rejected alternative. These ideas are central in regret theory (Bell, 1982; Loomes & Sugden, 1982), and the premise on which we built in this article.

Regret can be understood as an emotion of unrealized possibility (Kahneman in Lewis, 2016). Because unrealized possibilities are painful, people are generally regret averse, meaning that they dislike regret and try to avoid it in the future. Ample research had found that when choosing between gambles, decision makers take into account potential regret when they expect beforehand that they will learn the outcome of the unchosen gamble (e.g., Ritov, 1996; Zeelenberg, Beattie, Van der Pligt, & De Vries, 1996). Zeelenberg et al., for example, found that guaranteeing to always resolve one of two equally attractive gambles, causes more people to choose the gamble that will be resolved, because it rules out post-decisional comparison. They asked participants to choose between a safe gamble (65% chance of a smaller reward) and a risky gamble (35% chance of a larger reward). In one experimental manipulation, the experimenters informed participants that the risky gamble would always be resolved, irrespective of their decision. If participants would choose the safe gamble, they could later compare the outcome of their decision to the outcome of the rejected risky option.
The promise of post-decisional, counterfactual feedback on an uncertain outcome increased its desirability due to its regret-minimizing characteristics. That is, more people chose it because it precluded post-decisional comparison of outcomes (Zeelenberg et al., 1996): there was no possibility that the decision-maker would later find out that she should have chosen differently.

The effect of counterfactual feedback on decision-making under uncertainty has not only been studied in the context of choices between gambles, but also in various more real-life decision-contexts, such as lottery participation (Zeelenberg & Pieters, 2004), gym attendance (Van der Swaluw et al., 2018a), interpersonal bargaining (Zeelenberg & Beattie, 1997), or job negotiations (Larrick & Boles, 1995). Importantly, people in these experiments chose *now* for consequences in the nearby future. However, in none of these experiments, the exact delay between choice and outcome was made explicit or varied experimentally. Therefore, it remains unclear to what extent *future* feedback influences current decisions with future consequences. Although time is of key influence

in decision-making (Frederick, Loewenstein, & O'Donoghue, 2002), the role of counterfactual feedback in intertemporal choice remains unexplored. This is the topic of this article.

Intertemporal choice

Many decisions in life require trade-offs between costs and benefits that occur at different points in time (Loewenstein & Elster, 1992). For example, in order to have a good pension later in life, one must forgo spending at a younger age. In these intertemporal decisions, people generally prefer benefits sooner rather than later. Hence, the time until an outcome is obtained contributes to the desirability of that outcome (Loewenstein & Thaler, 1989). Ample research in economics and psychology has shown that waiting decreases desirability: delayed outcomes are valued less (i.e., discounted) than immediate outcomes (Ainslie, 1975; Samuelson, 1937).

Discounting of delayed outcomes generally does not happen in a constant fashion (e.g., 5% discounting per day), but is relatively stronger for sooner delay periods than for more distant delay periods (Kirby & Herrnstein, 1995; Mazur, 1987). Accordingly, people are described to be present biased: we disproportionally value benefits in the present over benefits in the future and choose impatiently (Laibson, 1997; O'Donoghue & Rabin, 1999).

Present bias has been linked to suboptimal decision-making, such as overeating (Ikeda, Kang, & Ohtake, 2010), insufficient physical activity (DellaVigna & Malmendier, 2006), smoking (Harrison, Lau, & Rutström, 2010), overspending and undersaving (Laibson, 1997; Thaler & Benartzi, 2004). As such, it is of theoretical and practical relevance to explore if counterfactual feedback can influence current decisions with future consequences.

Counterfactual feedback in intertemporal choice

Imagine a decision-maker who prefers an immediate and certain €25 over a 50% chance of €100 in one week, counterfactual feedback included (i.e., she always finds out the outcome of the gamble, also when choosing the certain €25). Accepting the certain €25 comes with a 50% chance of learning, one week later, that she should have chosen otherwise. As described above, Zeelenberg et al. (1996) have found clear feedback effects for gambles that were both uncertain (but differed in risk) and immediate. In a choice between two equally attractive options in the present, most participants chose the option that precluded post-decisional comparison of outcomes and thus prevented regret.

These effects of anticipated regret on behavior may be independent of timing of the negative consequences (Brewer, DeFrank, & Gilkey, 2016; Van der Swaluw et al., 2018b; Zeelenberg & Beattie, 1997). This can help explain the success of interventions that show that, if people are warned that they will learn the future consequences of current decisions, this can increase their number of gym visits (Van der Swaluw et al., 2018a), healthy food choices, or promote uptake of vaccinations (Koch, 2014).

Together, the literature suggests that the awareness of future feedback may influence decisions in the present. Hence, feedback is likely to influence intertemporal decision-making. We will examine this in three experiments in which we have participants choose between a certain outcome now and a risky gamble in the future. We expect that when there will be counterfactual feedback; the delayed outcome (the risky gamble) becomes more attractive because participants will want to prevent regret from finding out that the outcome of the gamble was better. More specifically, by giving the participants various choices between different certain outcomes and the future gamble, we can elicit the certainty equivalent of that future gamble. We expect these certainty equivalents to be higher when the participants expect feedback on the outcome of that gamble (Hypothesis 1).

If people value the future more and the present less as a result of counterfactual feedback, this may also result in less delay discounting (less impatience). In studies where people were asked to focus on their feelings after a decision (e.g., not exercising; Sandberg & Conner, 2008), or not using a condom (Richard, Van der Pligt & De Vries, 1996), they indicated more regret than when they were asked to focus on their feelings during the decision. As a consequence, people in these studies were more sensitive to the long-term consequences of their current decisions and behaved less impatient (exercised more and used condoms more often). In a similar vein, we expect that in the present experiment, when people know that there will be counterfactual feedback, they will discount the future less.

More specifically, we expect that when a participant chooses between a sure thing now and gambles at increasing delays, the gambles will be preferred less when they are further away in time (Blackburn & El-Deredy, 2013; Vanderveldt, Green, & Myerson, 2015). Subsequently, when we elicit the certainty equivalents of gambles at increasing delays, we expect that certainty equivalents will decline less when the participants know there will be feedback on the outcome of the gamble compared to when they know there will not be feedback (Hypothesis 2).

The present research

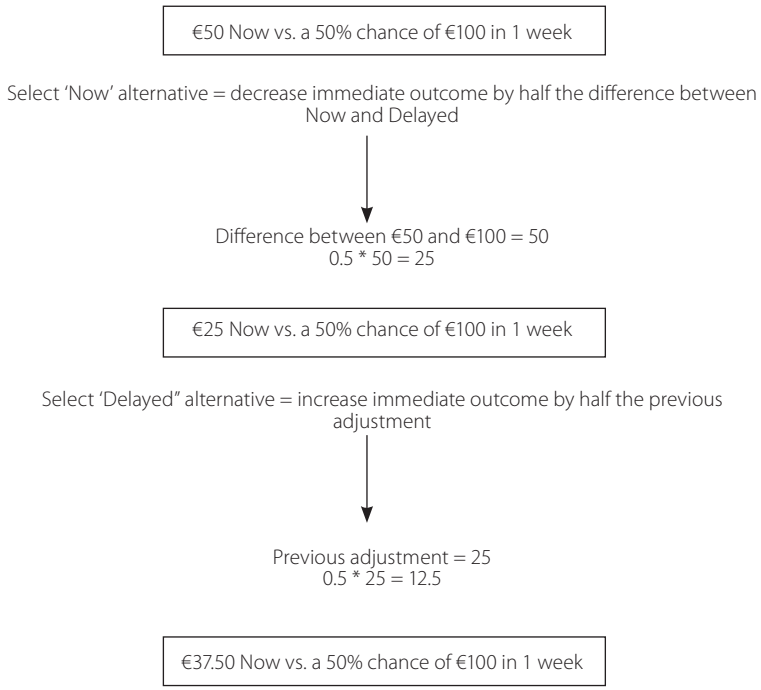
Below we discuss three experiments in which we manipulated whether counterfactual feedback on delayed gambles was present or not (Feedback Conditions vs. No Feedback Conditions), to study its effect on intertemporal decision-making. In all experiments, we used an amount-adjustment procedure that is commonly used to elicit delay discounting patterns (Blackburn & El-Deredy, 2013; Du, Green, & Myerson, 2002). Participants were presented with a series of forced choices between two hypothetical outcome amounts. One of these outcomes was immediate and certain (A) and the other was uncertain and delayed (B).

Participants were asked which of two decision-options they would prefer and, based on their response, the amount of the immediate option A was adjusted upwards or downwards in the next question for a total 4 times, using the half-the-difference algorithm (see Figure 1). This

procedure was designed to converge on an indifference point, where the certain immediate option is estimated to be subjectively equivalent to the gamble (i.e., the certainty equivalent).

In the Feedback Conditions, participants were informed that the gamble would always be resolved; that they would always get to know the outcome of the gamble and that choosing the certain option meant that the participant could compare this outcome to what would have happened if he or she had chosen otherwise. Every participant answered a total of five questions per six delays of the gamble (0, 1 week, 3 months, 6 months, 1 year, 5 years), resulting in six certainty equivalents. The certainty equivalents were our dependent variable.

Figure 1. Half the difference algorithm (based on Blackburn & El-Deredy, 2013; Du, Green, & Myerson, 2002).



The certain amount was always offered now and the delay in the Feedback Conditions entailed that participants would always learn the outcome of the gamble that took place now (0), after 1 week, 3 months, 6 months, 1 year, or 5 years. Therefore, choosing the immediate and certain amount meant that a participant could (hypothetically) compare these earnings to the outcome of the gamble only after the delay. In the No Feedback Condition, participants were informed that they would only learn the outcome of the delayed gamble if they would choose the gamble over the certain amount, and not otherwise. Therefore, choosing the immediate and certain amount meant that a participant could not compare the outcome to that of the gamble.

In Experiment 1, we tested our hypotheses in an online panel. Experiment 2, was a replication and extension the findings of Experiment 1 in a different and larger online sample. In Experiment 3, we incentivized the same procedure in a lab-setting, to test the hypotheses when the stakes were real. Below we describe these experiments in detail.

EXPERIMENT 1

In Experiment 1 we tested the hypotheses that the presence of feedback I) would result in higher certainty equivalents of delayed gambles, and II) would attenuate delay discounting of uncertain outcomes. We used two different amounts as outcome of the gambles (€100 vs. €5000) to account for a potential magnitude effect in delay discounting (Green, Myerson, & McFadden, 1997; Thaler, 1981). The experiment was reviewed by the Ethics Review Board (EC-2017.EX71) and all participants signed informed consent.

Method

Participants and design

Panel members of the ISO-certified CG-panel were recruited to adequately represent the Dutch population and participated in an online survey in return for a financial compensation by CG Insights, (N = 207, 102 (49.5%) female; M_{age} = 51.32 years, SD = 15.11). Participants were randomly assigned to one the conditions of the 2 (No Feedback vs. Feedback) × 2 (50% chance of €100 vs. 50% chance of €5000) design.

Procedure and materials

In order to explain to the participants the type of decisions that they were going to make, the experiment started with a graphic explanation of Joe, who is asked to choose between a certain reward and a 50% chance of €100 [€5000] to be determined by a coin flip, both immediately available. Joe picks the certain reward and gets paid. Participants were informed that the coin would have been flipped if Joe would have picked the gamble, but not otherwise.

In the Feedback Condition, participants watched an additional graphic description of Rob. Rob's decision-options were quantitatively identical to Joe's, but the coin flip would always take place, regardless of Rob's decision. The scenario pictured Rob choosing the certain reward and witnessing the coin flip being resolved in his disadvantage. That is, if Rob had chosen the gamble, he would have earned more. In the Feedback Condition, participants next answered the following question: "Joe and Rob made the exact same amount of money. Who do you think is happiest? A) Joe, B) Rob, C) equally happy."

Next, participants completed the forced choice task (see above). As described above, each participant was asked to select his or her preference between two options five times. In repeated sequences, the gamble would take place at six different time periods (now (0), after 1 week, 3 months, 6 months, 1 year, or 5 years). As such, each participant made a total of 30 hypothetical choices, resulting in 6 certainty equivalents.

We predicted that the certainty equivalents would be higher in the Feedback Condition than in the No Feedback Condition (H1) and that certainty equivalents would decline less in the Feedback Condition than in the No Feedback Condition (H2).

Results

The certainty equivalent represents the desirability of the (delayed) gamble. Across 6 points in time, certainty equivalents for each period can be plotted to present a discounting curve (see Figures 2 and 3). The certainty equivalent with delay 0 is the valuation of the gamble in the present, which was used as the starting point of the discounting curve (Blackburn & El-Deredy, 2013).

Of the 207 participants, 31 were excluded due to incomplete data or random responding. Analyses including all participants to test for robustness yielded similar results. Based on the algorithm for identifying nonsystematic discounting by Johnson and Bickel (2008), 23 participants showed nonsystematic discounting on 1 certainty equivalent (e.g., 30-28-23-17-**68**-10). The diverging certainty equivalents were equalized to their predecessor and these participants remained included in the dataset (e.g., 30-28-23-17-**17**-10).

Within the Feedback Condition, 84% of participants stated that Joe felt happier than Bob, despite the fact that both earned the same amount. This suggests that the majority of participants understood the consequences of counterfactual feedback in the current experimental context.

Feedback effects over time

Table 1 displays the descriptive statistics per condition and differences between conditions. Figures 2 and 3 display the discounting curves with and without feedback for the €100 and €5000 condition respectively. Because the distribution of certainty equivalents was skewed, as is typically the case (Myerson, Green, & Warusawitharana, 2001), and because log-transformations did not normalize the data, we used the non-parametric Mann-Whitney test (see: Siegel & Castellan, 1988) to test the hypothesis that the certainty equivalents in the Feedback Conditions would be higher than in the No Feedback Conditions (H1).

Table 1. Certainty equivalents of gambles in the two Feedback conditions of Experiment 1

	No Feedback	Feedback	Difference
50% chance of €100	(*N* = 48)	(*N* = 42)	
Mean *(SD)*	18.90 (12.37)[a]	25.05 (15.16)	6.15
Median	16.40	22.66	6.26
Range	1.56 - 52.08	5.73 - 70.83	
50% chance of €5000	(*N* = 47)	(*N* = 38)	
Mean *(SD)*	476.51 (504.12)	601.70 (693.53)	125.19
Median	260.42	507.81	247.39
Range	78.13 - 2421.88	78.13 - 3046.88	

[a] Numbers reflect aggregated (over time) certainty equivalents of gambles per condition in €.

Figure 2. Certainty equivalents of a 50% chance of €100 with and without feedback in Experiment 1

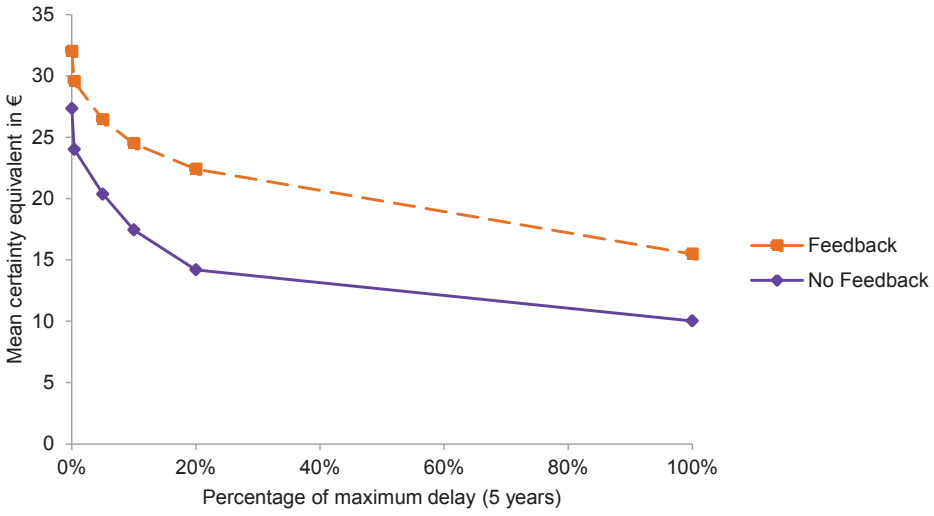


Figure 3. Certainty equivalents of a 50% chance of €5000 with and without feedback in Experiment 1

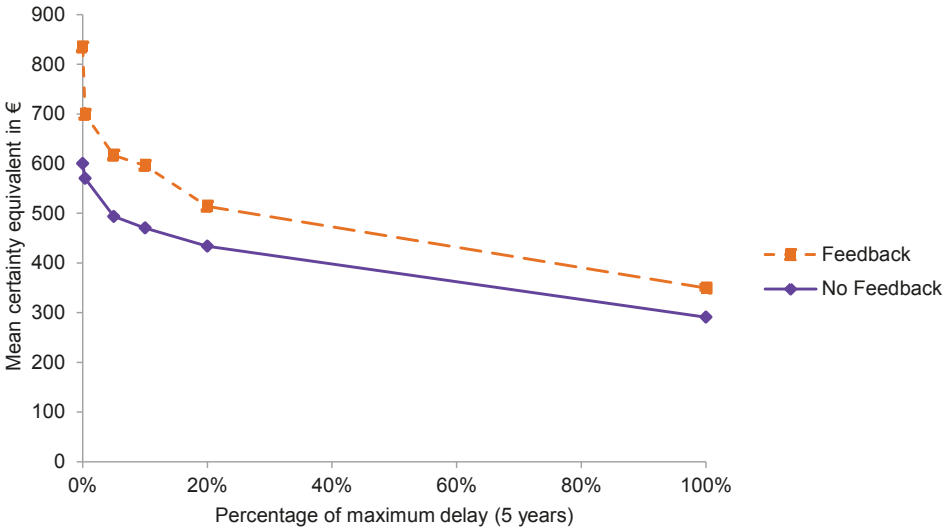


Table 2 displays the outcomes of the Mean Rank comparisons of the average certainty equivalent and per delay period separately. As can be seen in Figures 2 and 3, certainty equivalents of (delayed) gambles were structurally higher in the Feedback Condition than in the No Feedback Condition. Expressed in proportions of expected value, feedback increased the average certainty equivalent of the €100-gamble by 12% ((25.05-18.90) /50) and the €5000-gamble by 5% ((601.70-476.51)/ 2500). The difference in average certainty equivalent was statistically significant in the €100-condition (Median $_{\text{Feedback}}$ = €22.66 and Median $_{\text{No Feedback}}$ = €16.40, U = 787.50, p (one-tailed) = .04), but not in the €5000-condition (Median $_{\text{Feedback}}$ = €507.81, Median $_{\text{No Feedback}}$ = €260.42, U = 792.50, p (one-tailed) = .18).

Table 2. Mean ranks of certainty equivalents & Area Under the Curve in Experiment 1

	Mean Rank of certainty equivalents [a]							Mean Rank AUC
50% chance of €100	now	1 week	3 months	6 months	1 year	5 years	Average [b]	
No Feedback (*N* = 48)	42.55	41.36	40.84	39.56	38.20	40.84	40.91	43.10
Feedback (*N* = 42)	48.87	50.23	50.82	52.29	53.85	50.82	50.75	48.24
Mann-Whitney *U*	866.50	809.50	784.50*	723.00*	657.50**	784.50*	787.50*	893.00
50% chance of €5000								
No Feedback (*N* = 47)	40.72	41.27	40.93	41.09	41.67	42.65	40.86	45.14
Feedback (*N*= 38)	48.84	45.14	45.57	45.37	44.64	43.43	45.64	40.36
Mann-Whitney *U*	786.00	811.50	795.50	803.00	830.50	876.50	792.50	792.50

[a] The mean of rank scores per condition
[b] The mean of rank scores of the average certainty equivalent over the six time periods
AUC = Area Under the empirical discounting Curve
* significant at *p* (one-tailed) = .05
** significant at *p* (one-tailed) = .01

The effect of feedback on delay discounting

To test the hypothesis that certainty equivalents would decline less as function of delay in the Feedback Condition than in the No Feedback Condition (H2), the empirical Area Under the discounting Curve (AUC) was calculated for each participant. The AUC is an expression of the area under the curve as a proportion (between 0 and 1) of the maximum possible area. Accordingly, steeper discounting results in a lower AUC. Because the AUC is calculated from the actual data points, rather than from a data-fitting procedure, no theoretical assumptions on the form of the discounting curves (e.g., exponential or (quasi)- hyperbolic) had to be made (Myerson et al., 2001).

The AUC distribution was skewed and, therefore, also analyzed with Mann-Whitney tests. First, we tested whether levels of discounting differed between amounts (the magnitude effect). Discounting was stronger in the €100-condition (Median AUC = .63) than in the €5000 condition (Median AUC = .73, U = 3065.50, *p* = .02). Hence, differences in discounting between Feedback Conditions were also analyzed separately per amount.

To test differences in AUC between Feedback Conditions, two independent (€100 and €5000-condition) Mann-Whitney tests were performed. In the €100-condition, Feedback (Median AUC = .66) resulted in moderately less discounting than No Feedback (Median AUC = .52), but this was not statistically significant (U = 893.00, *p* (one-tailed) = .13). In the €5000-condition, we also observe no significant differences in discounting between the Feedback Condition (Median AUC = .71) and the No Feedback Condition (Median AUC = .76, U = 792.50, *p* (one-tailed) = .18).

Discussion

Experiment 1 demonstrates that the presence of future feedback influences present decisions with delayed consequences. In a choice between an immediate certain amount and a delayed 50% chance of €100, people in the Feedback Condition value the delayed and uncertain outcome higher than people in the No Feedback Condition. This supports Hypothesis 1 that feedback increases certainty equivalents of delayed uncertain outcomes. This pattern was the same in the €5000 condition, but not statistically significant. Furthermore, we find no compelling evidence for Hypothesis 2 that feedback would attenuate delay discounting.

EXPERIMENT 2

In Experiment 2, we explored the generalizability of our findings in a different and larger sample and with one different amount. Given the magnitude effect in Experiment 1, we still distinguished a low and high payoff, but changed the high amount from €5000 to €1000. The €100-condition remained identical. This setting allowed us to test the robustness of our findings in a near to similar experimental context. Experiment 2 was also reviewed by the Ethics Review Board (EC-2017.EX71at) and all participants signed informed consent.

Method

Apart from the change in payoff from €5000 to €1000, every other aspect in the task of Experiment 2 was identical to Experiment 1. Panel members of the ISO-certified Motivaction-panel were recruited to adequately represent the Dutch population and participated in return for a financial compensation by Motivaction ($N = 735$, 323 (43.9%) female; $M_{age} = 45.30$ years, $SD = 15.00$). Participants were randomly assigned to one of four conditions of the 2 (No Feedback vs. Feedback) × 2 (50% chance of €100 vs. €1000) between-subjects design.

Results

Of the 738 participants, 146 were excluded due to incomplete data or random responding. Similar to Experiment 1, analyses including all participants to test for robustness yielded similar results. Based on the algorithm for identifying nonsystematic discounting by Johnson and Bickel (2008), 88 participants showed nonsystematic discounting on 1 indifference point (e.g., 30-28-23-17-**68**-10). As in Experiment 1, the diverging indifference points were equalized to their predecessor and these participants remained included in the dataset (e.g. 30-28-23-17-**17**-10).

Within the Feedback Condition, 83% of participants stated that Joe felt happier than Bob, despite both earning the same amount. This indicates that the majority of participants understood the consequences of counterfactual feedback in the experiment.

Feedback effects over time

Table 3 displays the descriptive statistics per condition and Figures 4 and 5 display the discounting curves with and without feedback in the €100 and €1000 condition respectively.

Table 3. Certainty equivalents of gambles without and with feedback in Experiment 2

	No Feedback	Feedback	Difference
50% chance on €100	(*N* = 169)	(*N* = 147)	
Mean (SD)	19.70 (15.01) [a]	22.95 (15.39)	3.25
Median	16.67	20.31	3.64
Range	1.56 - 96.88	1.56 - 82.29	
50% chance on €1000	(*N* = 136)	(*N* = 137)	
Mean (SD)	159.39 (140.28)	168.68 (152.75)	9.29
Median	112.00	109.38	-2.62
Range	15.63 - 619.79	15.63 - 822.92	

[a] Numbers reflect aggregated (over time) certainty equivalents of gambles per condition in €.

Figure 4. Certainty equivalents of a 50% chance of €100 with and without feedback in Experiment 2

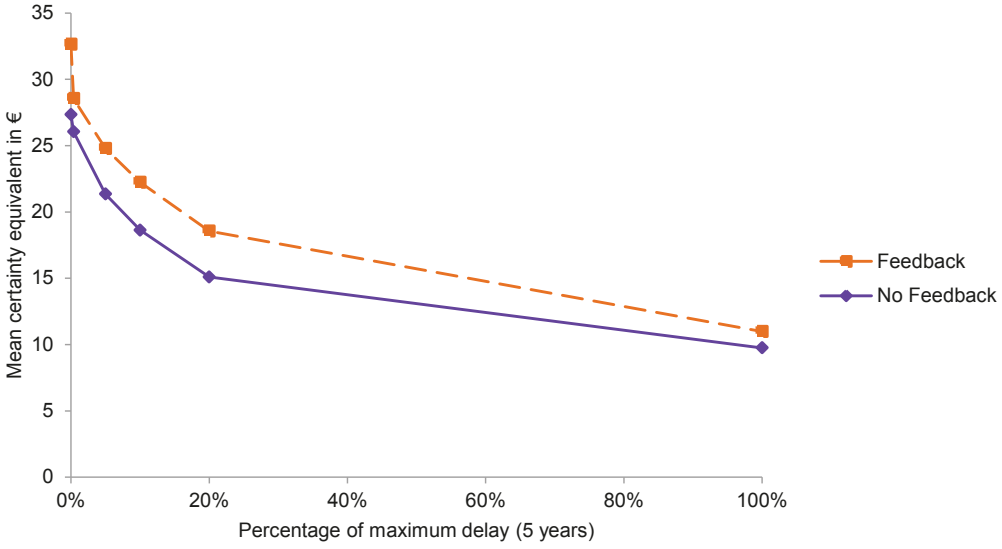


Figure 5. Certainty equivalents of a 50% chance of €1000 with and without feedback in Experiment 2

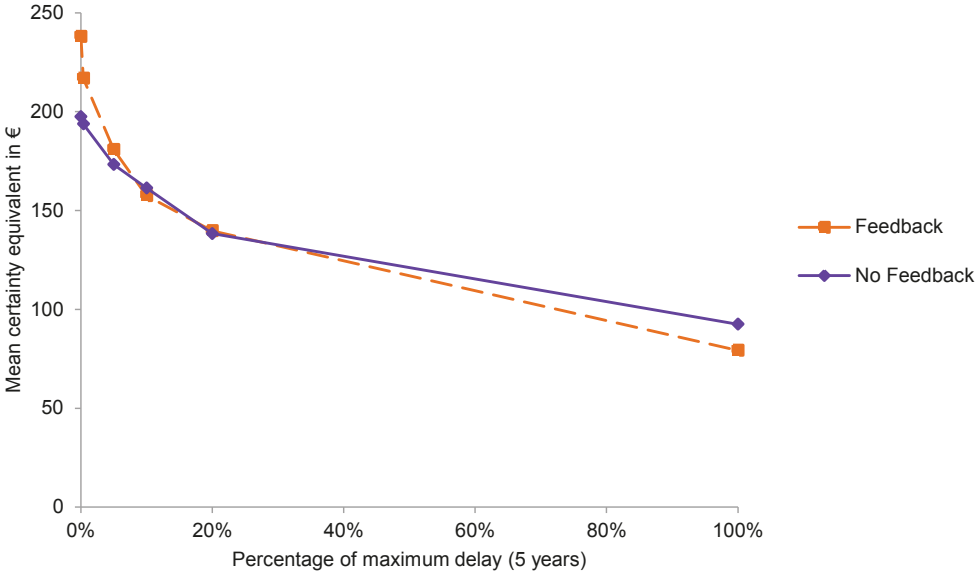


Table 4 displays the outcomes of the Mean Rank comparisons of the average certainty equivalent and per delay period separately. Similar to Experiment 1, certainty equivalents in the €100-condition were all higher in the Feedback Condition than in the No Feedback Condition. Expressed as a proportion of the expected value, feedback increased the average certainty equivalent of the €100-gamble by 6.5% and the €1000-gamble by 1.86%. In the €100-condition, the difference in average certainty equivalent was statistically significant (Median $_{\text{Feedback}}$ = €20.31 and Median $_{\text{No Feedback}}$ = €16.67, U = 10694.00, p (one-tailed) = .02), but not in the €1000-condition (Median $_{\text{Feedback}}$ = €109.38, Median $_{\text{No Feedback}}$ = €112.00, U = 8899.00, p (one-tailed) = .26).

Table 4. Mean ranks of certainty equivalents & Area Under the Curve in Experiment 2

	Mean Rank of certainty equivalents[a]							Mean Rank AUC
50% chance of €100	now	1 week	3 months	6 months	1 year	5 years	Average[b]	
No Feedback (*N* = 169)	147.43	152.23	149.19	148.22	146.92	152.63	148.28	155.19
Feedback (*N* = 147)	171.22	165.70	169.20	170.31	171.81	165.25	170.25	162.31
Mann-Whitney *U*	10551.00*	11362.50	10848.50*	10685.00*	10465.00**	11429.00	10694.00*	11862.00
50% chance of €1000								
No Feedback (*N* = 136)	129.31	132.35	136.36	137.63	136.01	137.83	133.93	148.83
Feedback (*N* = 137)	144.64	141.62	137.63	136.37	137.99	136.18	140.04	125.26**
Mann-Whitney *U*	8270.00	8683.00	9229.50	9230.00	9181.00	9203.00	8899.00	7707.00

[a] The mean of rank scores per condition
[b] The mean of rank scores of the average certainty equivalent over the six time periods
AUC = Area Under the empirical discounting Curve
* significant at *p* (one-tailed) = .05
** significant at *p* (one-tailed) = .01

The effect of feedback on delay discounting

To identify if levels of discounting differed between amounts (the magnitude effect), Mann-Whitney tests with AUC as the dependent variable and amount-condition (€100 vs. €1000) as the independent variable were performed. As in Experiment 1, the lower amount was discounted stronger (Median $AUC_{€100}$ = .47) than the higher amount (Median $AUC_{€1000}$ = .66, U = 34713.00, $p < .001$). Accordingly, differences in discounting between Feedback Conditions were also analyzed separately per amount.

To test differences in discounting between feedback-conditions, two independent (€100 and €1000 condition) Mann-Whitney tests were performed with AUC as the dependent variable and Feedback vs. No Feedback as the independent variable. Similar to Experiment 1, in the €100-condition, discounting with feedback (Median AUC = .47) did not differ significantly from discounting without feedback (Median AUC = .48, U = 11862.00, *p* (one-tailed) = .25). In the €1000-condition, we observe a statistically significant feedback effect that was contrary to our hypothesis (H2). Discounting was stronger in the Feedback Condition (Median AUC = .60) than in the No Feedback Condition (Median AUC = .73, U = 792.50, $p = .01$).

Discussion

Experiment 2 replicates several results of Experiment 1 in a different and larger sample. The certainty equivalent of a delayed 50% chance of €100 increased by the presence of counterfactual feedback. This pattern was nearly similar in the €1000-condition, but not statistically significant.

Similar to Experiment 1, the higher amount was discounted stronger than the lower amount. In the €1000-condition, feedback significantly increased delay discounting. As before, feedback can influence delayed uncertain outcomes, but we find no evidence for feedback attenuating delay discounting (H2).

EXPERIMENT 3

Experiments 1 and 2 revealed that counterfactual feedback increases the desirability of a delayed and uncertain €100, compared to an immediate sure gain. The choices in these experiments were non-consequential, as the experiments were not incentivized. The interpretation of the results relies on the assumption that participants can imagine the situation and have no reason to disguise their true preferences (Kahneman & Tversky, 1979). Therefore, in Experiment 3 we tested hypothesis 1, that promising feedback would result in higher certainty equivalents of delayed uncertain outcomes, using an incentivized procedure in a lab. For practical reasons, we chose to use only the €100 gamble and we limited the delay to 1 week. Experiment 3 was reviewed by the Ethics Review Board (EC-2017.EX71a) and all participants signed informed consent

Method

Participants and design

Tilburg University students participated in a double blinded experiment in return for course credit or a show-up fee ($N = 346$, 231 (66.80%) female; $M_{age} = 20.07$ years, $SD = 3.22$). Participants were randomly assigned to one of two conditions (No Feedback vs. Feedback). Upon enrolling for the experiment, participants were informed that participation always required two visits that were one week apart. This way, preferences for immediate payment could not be attributed to participant's unwillingness to come back a week later (i.e., transaction costs).

Procedure

Participants were seated in a separate cubicle to fill out the survey on a computer. The task was similar to Experiments 1 and 2, but limited to the €100-amount and a delay of 1 week. After reading about Joe (No Feedback) and Rob (Feedback), participants in the Feedback Condition answered who they thought felt happiest. Next, participants

in both conditions stated their preferences between two amounts across two time periods: now and 1 week. The immediate certain option started at €50 and varied based on preferences according to the algorithm also used in Experiments 1 and 2 (Figure 1).

We incentivized the task by informing participants prior to the task that one in every 10 participants would be selected for actual payment after the task (cf. Johnson and Bickel, 2002). We emphasized that this implied that every decision could potentially determine their earnings. In the Feedback Condition, participants were informed that, if they would be selected, we would *always* flip the coin in their presence regardless of their decision (counterfactual feedback). The whole procedure was explained graphically to all participants and the instructions could be repeated if participants indicated that they did not understand it. All participants stated that they understood the procedure.

Results

Of the 346 participants, 6 were excluded due to errors in the experimental procedure. As in Experiments 1 and 2, this did not affect the pattern of results. Of the 173 participants in the Feedback Condition, 161 (93%) stated that Joe felt happier than Rob, again indicating understanding of the emotional consequences of the feedback.

Table 5 displays the descriptive statistics per Feedback Condition. Table 6 displays the outcomes of the Mean Rank comparisons of the average certainty equivalent and per delay period separately. Similar to Experiments 1 and 2, certainty equivalents in the €100-condition were structurally higher in the Feedback Condition than in the No Feedback Condition. Expressed as a proportion of the expected value, feedback increased the average certainty equivalent of the €100-gamble by 8.4%. The difference was statistically significant (Median $_{\text{Feedback}}$ = €29.69 and Median $_{\text{No Feedback}}$ = €21.88, U = 25898.00, p (one-tailed) < .01).

Table 5. Certainty equivalents of gambles without and with feedback in Experiment 3

	No Feedback	Feedback	Difference
50% chance of €100	(*N* = 167)	(*N* = 173)	
Mean (SD)	25.26 (11.69)[a]	29.81 (14.29)	4.55
Median	21.88	29.69	7.81
Range	1.56 - 60.94	4.69 - 98.44	

[a] Numbers reflect aggregated (over time) certainty equivalents of gambles per condition in €.

Furthermore, feedback increased the certainty equivalent of the gamble when there was no delay (Median $_{\text{Feedback}}$ = €29.69 and Median $_{\text{No Feedback}}$ = €23.44, U = 11917.50, p (one-tailed) < .01) and when feedback was guaranteed after a week (Median $_{\text{Feedback}}$ = €29.69 and Median $_{\text{No Feedback}}$ = €20.31, U = 11917.50, p (one-tailed) < .01). As in the previous experiments, the feedback manipulation did not influence the degree of discounting.

Table 6. Mean ranks of certainty equivalents & Area Under the Curve in Experiment 3

	Mean Rank of certainty equivalents[a]			Mean Rank AUC
50% chance of €100	now	1 week	Average[b]	
No Feedback (*N* = 167)	155.36	155.36	155.08	168.51
Feedback (*N* = 173)	185.11	185.11	185.39	172.42
Mann-Whitney *U*	11917.50**	11917.50**	11870.00**	14113.00

[a] The mean of rank scores per condition
[b] The mean of rank scores of the average certainty equivalent over the two time periods
AUC = Area Under the empirical discounting Curve
** significant at *p* (one-tailed) = .01

Discussion

Experiment 3 replicates results from Experiments 1 and 2 in an incentivized lab setting. The certainty equivalent of a 50% chance of €100 increased when there was counterfactual feedback. This effect remained when the feedback was promised after 1 week. With a real delay and with real money at stake, participants were interested in preventing regret in the future and chose accordingly in the present. Feedback did not affect discounting.

GENERAL DISCUSSION

In three experiments we manipulated counterfactual feedback on delayed uncertain outcomes to study regret aversion in intertemporal choice. Across all experiments, feedback did not attenuate delay discounting. Delayed feedback, however, did influence present decisions with delayed consequences in all three experiments. The knowledge that a gamble for €100 would always be resolved increased the desirability of this gamble compared to an immediate sure gain. With feedback, the choice to gamble instead of accepting a certain payment reduced the chances of having to cope with the emotional consequences of a forgone and unrealized possibility: regret. This replicates and extends previous research in a novel intertemporal context.

Previous studies have found that counterfactual feedback on high-risk gambles can promote risk seeking in a choice between a low-risk gamble and a higher-risk gamble (Ritov, 1996; Zeelenberg et al., 1996). The authors explain their findings as resulting from anticipated regret; choosing the gamble that was guaranteed to be resolved precluded the possibility of post-decisional comparison. The results of the present study support this reasoning, replicate previous findings and further suggest that feedback elevates the desirability of uncertain outcomes if the alternative is not just lower risk but even complete certainty. Most important, the results of the present experiments demonstrate that feedback maintains its significant influence when the certain payment is immediate and the (resolution of the) uncertain option is delayed.

Feedback did not significantly influence the certainty equivalent of high-stakes gambles (€1000 and €5000). This might be explained by the relatively high certain amount that decision-makers have to reject in order to gamble (Zeelenberg, 1999). In our experimental design, choosing the gamble in the Feedback Conditions did not entirely shield the decision-maker from regret. When choosing the gamble, there was also a 50% chance for regret (winning €0, while missing out on the sure gain). Taking a risk and subsequently losing comes with a vivid counterfactual and possibly regret; "I should have chosen the certain amount".[1] In the €100-conditions, this loss is -in absolute terms- lower than in the €5000 or €1000 conditions. Although the potential gains of gambling in the high-amount conditions were much higher, so were the potential losses. People generally respond stronger to losses than to gains of equal magnitude and whether something is perceived as a loss or a gain depends on their point of reference (Kahneman & Tversky, 1979). If a decision-maker accepts a certain payment and finds out she would have earned more if she had gambled, this may cause regret over *forgone gains*. If a decision-maker rejects a certain gain and eventually loses the gamble, this may cause regret over *losing* the certain gain. Logically, this loss aversion is higher in the €5000 and €1000 conditions than in the €100-conditions.

In future research, it would be valuable to explore which reference points participants use in the comparison of reality to the alternative reality. If the outcome of the rejected certain gain serves as the reference point (Ritov, 1996), anticipated regret over forgone gains and losses may cancel each other out and attenuate feedback effects when high-stakes gambles oppose high certain amounts.

Feedback did not attenuate delay discounting across all experiments. Despite the observation that feedback increased the certainty equivalent of delayed uncertain outcomes (H1), these gambles were mostly discounted at similar rates as gambles without feedback. This might be caused by a specific element of our experimental design. Participants in the Feedback Conditions were promised feedback at delay 0 and subsequently after 1 week, 3 months, 6 months, 1 year and 5 years. In repeated sequences, the only feature that we varied was time. This may have caused participants to only incorporate the (additional) delay in their later decisions, while already accounting for the threat of feedback in their first sequence (delay 0).

This reasoning can be reconciled with discounting research by Blackburn and El-Deredy (2013); Cox and Dallery (2016), which showed that adding additional decision-features to a delayed

[1] In fact, in developing Prospect Theory, Kahneman and Tversky (1979) first considered regret as the most important cause for why people are risk averse in the domain of gains (Lewis, 2016). They later abandoned this theorizing as the primary cause of their findings, which resulted in arguably the most accurate description of decision-making under uncertainty. Later, Kahneman (2011) stated that a blind spot of Prospect Theory is that the theory does not allow for counterfactual thinking, regret and disappointment.

reward (i.e., uncertainty about amount or probability) can cause changes in preferences that can be easily modelled, if participants' responses to those features in the present (delay 0) are taken into account. In future studies, feedback might affect discounting if starting points (delay 0) of experimental conditions are identical, and the Feedback Condition incorporates feedback only when the gamble is delayed.

In Experiment 2, feedback unexpectedly increased discounting of the €1000-gamble. Although average certainty equivalents of this gamble were higher in the Feedback Condition up to 3-months delay (though, not-significantly), feedback caused participants to discount this gamble stronger. Despite nonsignificant differences in certainty equivalent between Feedback Conditions, it might be that feedback nonetheless had a non-negligible impact on the value of the €1000-gamble when it was immediate (the difference between Feedback Conditions at delay 0 in Figure 5 is €40.60), and that feedback eventually lost most of its (initially small) impact over time, causing a steeper decline of certainty equivalents with additional delay.

If this reasoning fits empirical observations in future experiments, this may indicate that for feedback to uphold its impact on decision-making, it needs additional reinforcement (e.g., reminders or social impact; Zeelenberg & Pieters, 2007) when it was initially guaranteed immediately and subsequently promised after a delay. It would be valuable to explore the discounting pattern if starting conditions of between-subjects groups (at delay 0) are identical and the Feedback-condition guarantees feedback only when the gamble is delayed.

This study was the first systematic exploration of feedback in intertemporal choice and was necessarily limited in multiple ways. First, previous research has shown interpersonal differences in regret tendency and according responses to counterfactual feedback (Schwartz et al., 2002; Van der Swaluw et al., 2018b). These differences may also influence feedback effects in intertemporal decision-making and could be accounted for in a within-subjects experimental design (participants than serve as their own control). To avoid cross-over effects, participant fatigue or decreases in commitment to the experiment, we did not vary the different amounts or the feedback-variations within subjects. Doing so would have resulted in 120 choices per participant in contrast to the more tolerable 30 choices. Nonetheless, it would be valuable to explore methods to vary the feedback manipulations within subjects or add measures of regret tendency (e.g., the Regret Scale; Schwartz et al., 2002) and demographics to the procedure, to explore mediating or moderating factors.

Second, we used a probability of 50% in all studies. Modelling studies suggest that probability and delay both have a unique impact on- and interact in the determination of a certainty equivalent (Green, Myerson, & Ostaszewski, 1999; Vanderveldt et al., 2015). Therefore, it would be valuable to also explore the effects of feedback on delay discounting with variating probabilities. Third, as in inventive experimentation by Blackburn and El-Deredy (2013), the

immediate option in the current experiments was also the certain option. Traditionally, feedback effects are mostly studied in decisions between two uncertain decision-options (Larrick & Boles, 1995; Ritov, 1996; Zeelenberg et al., 1996). It would be of theoretical interest to also explore the feedback effects if both the immediate option and the delayed option are uncertain.

The results of the present experiment may also have practical implications. Thaler (1981) metaphorically explains present biased decision-making as a conflict between two 'selves'; a planner and a doer. The planner prefers to save for retirement, to eat fruit instead of chocolate or to exercise instead of relaxing, but when the future has 'arrived', the present biased doer takes over. He overrules previous preferences and yields to the immediate gratification of spending, unhealthy nutrition and sedentary behavior. Decisions with an intertemporal structure are key determinants of our health and well-being (Forouzanfar et al., 2016). Therefore, the present experiments might promote further development of instruments via which policy makers, employers or health professionals can provide counterfactual feedback on delayed consequences of current decisions (e.g., Patel et al., 2016; Van der Swaluw et al., 2018a). This may increase the desirability of the delayed outcome. Given the non-effect of feedback when amounts where higher, interventionist should be aware of the relation between the immediate pain and the future pleasure of the decision alternatives. If a decision-maker has to reject a too large immediate and certain benefit for an uncertain and delayed outcome, merely feedback might not alter decisions.

Conclusion

It has been long known that people alter their decisions upon expecting counterfactual feedback, preventing future regret. In general, people also prefer benefits sooner rather than later. It remained unexplored to what extent counterfactual feedback influenced intertemporal choice. The results of three experiments indicate that counterfactual feedback on a delayed uncertain outcome increases the desirability of the outcome. As an instrument, counterfactual feedback may help overcome societal and organizational issues that stem from intertemporal decision biases. When the uncertain outcome is higher (€5000 or €1000), so is the to-be rejected certain amount, which may diminish feedback effects. Feedback did not attenuate delay discounting and we recommend more advanced experimental designs to further explore this relation.

REFERENCES

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological Bulletin, 82*(4), 463.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research, 30*(5), 961-981.

Blackburn, M., & El-Deredy, W. (2013). The future is risky: Discounting of delayed and uncertain outcomes. *Behavioural Processes, 94*, 9-18.

Brewer, N. T., DeFrank, J. T., & Gilkey, M. B. (2016). Anticipated regret and health behavior: A meta-analysis. *Health Psychology, 35*(11), 1264-1275.

Cox, D. J., & Dallery, J. (2016). Effects of delay and probability combinations on discounting in humans. *Behavioural Processes, 131* (15), 15-23.

DellaVigna, S., & Malmendier, U. (2006). Paying not to go to the gym. *The American Economic Review, 96* (3), 694-719.

Du, W., Green, L., & Myerson, J. (2002). Cross-cultural comparisons of discounting delayed and probabilistic rewards. *The Psychological Record, 52*(4), 479.

Forouzanfar, M. H., Afshin, A., Alexander, L. T., Anderson, H. R., Bhutta, Z. A., Biryukov, S., . . . Charlson, F. J. (2016). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388*(10053), 1659.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature, 40*(2), 351-401.

Green, L., Myerson, J., & McFadden, E. (1997). Rate of temporal discounting decreases with amount of reward. *Memory & Cognition, 25*(5), 715-723.

Green, L., Myerson, J., & Ostaszewski, P. (1999). Amount of reward has opposite effects on the discounting of delayed and probabilistic outcomes. *Journal of Experimental Psychology-Learning Memory and Cognition, 25*(2), 418-427.

Harrison, G. W., Lau, M. I., & Rutström, E. E. (2010). Individual discount rates and smoking: Evidence from a field experiment in Denmark. *Journal of Health Economics, 29*(5), 708-717.

Ikeda, S., Kang, M.-I., & Ohtake, F. (2010). Hyperbolic discounting, the sign effect, and the body mass index. *Journal of Health Economics, 29*(2), 268-284.

Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the Experimental Analysis of Behavior, 77*(2), 129-146.

Johnson, M. W., & Bickel, W. K. (2008). An algorithm for identifying nonsystematic delay-discounting data. *Experimental and Clinical Psychopharmacology, 16*(3), 264 -274.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Penguin.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-292.

Kirby, K. N., & Herrnstein, R. J. (1995). Preference reversals due to myopic discounting of delayed reward. *Psychological Science, 6*(2), 83-89.

Koch, E. J. (2014). How does anticipated regret influence health and safety decisions? A literature review. *Basic and Applied Social Psychology, 36*(5), 397-412.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics, 112*(2), 443-477.

Larrick, R. P., & Boles, T. L. (1995). Avoiding Regret in Decisions with Feedback: A Negotiation Example. *Organizational Behavior and Human Decision Processes, 63*(1), 87-97.

Lewis, M. (2016). *The Undoing Project: A Friendship that Changed the World*. New York: W.W. Norton & Company.

Loewenstein, G., & Elster, J. (1992). *Choice over time*. New York: Russell Sage Foundation.

Loewenstein, G., & Thaler, R. H. (1989). Anomalies: intertemporal choice. *The Journal of Economic Perspectives, 3*(4), 181-193.

Loomes, G., & Sugden, R. (1982). Regret Theory: An Alternative Theory Of Rational Choice Under Uncertainty. *The Economic Journal*, *92*(368), 805-824.

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In Commons, M.L., Mazur, J.E., Nevin, J.A. (Eds), *Quantitative Analyses Of Behavior- The Effect of Delay and of Intervening Events on Reinforcement* (55-73). New Jersey: Lawrence Erlbaum Associates.

Myerson, J., Green, L., & Warusawitharana, M. (2001). Area under the curve as a measure of discounting. *Journal of the Experimental Analysis Of Behavior, 76*(2), 235-243.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, *89* (1), 103-124.

Patel, M. S., Asch, D. A., Rosin, R., Small, D. S., Bellamy, S. L., Eberbach, K., . . . Wesby, L. (2016). Individual versus team-based financial incentives to increase physical activity: a randomized, controlled trial. *Journal of General Internal Medicine*, *31*(7), 746-54.

Richard, R., Van der Pligt, J., & De Vries, N. (1996). Anticipated regret and time perspective: Changing sexual risk-taking behavior. *Journal of Behavioral Decision Making, 9*(3), 185-199.

Ritov, I. (1996). Probability of regret: Anticipation of uncertainty resolution in choice. *Organizational Behavior and Human Decision Processes, 66*(2), 228-236.

Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies*, *4*(2), 155-161.

Sandberg, T., & Conner, M. (2008). Anticipated regret as an additional predictor in the theory of planned behaviour: A meta-analysis. *British Journal of Social Psychology, 47*(4), 589-606.

Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: happiness is a matter of choice. *Journal of Personality and Social Psychology, 83*(5), 1178.

Siegel, S., & Castellan, N. J. (1988). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill Book Company.

Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics letters, 8*(3), 201-207.

Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of political Economy*, *112*(S1), S164-S187.

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Schipper, M., Zeelenberg, M., Berkhout, S., Polder, J.J., & Prast, H. M. (2018)a. Commitment lotteries promote physical activity among overweight adults—a cluster randomized trial. *Annals of Behavioral Medicine*, 52(4), 342-351.

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Zeelenberg, M., Polder, J. J., & Prast, H. M. (2018)b. Emotional responses to behavioral economic incentives for health behavior change. *Psychology, Health & Medicine (Online)*, 1-10.

Vanderveldt, A., Green, L., & Myerson, J. (2015). Discounting of monetary rewards that are both delayed and probabilistic: Delay and probability combine multiplicatively, not additively. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(1), 148.

Von Neumann, J., & Morgenstern, O. (1944, 2007). *Theory of games and economic behavior (60th Anniversary Commemorative Edition)*. New Jersey: Princeton university press.

Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making*, 12(2), 93-106.

Zeelenberg, M., & Beattie, J. (1997). Consequences of regret aversion 2: Additional evidence for effects of feedback on decision making. *Organizational Behavior and Human Decision Processes*, 72(1), 63-78.

Zeelenberg, M., Beattie, J., Van der Pligt, J., & De Vries, N. K. (1996). Consequences of regret aversion: Effects of expected feedback on risky decision making. *Organizational Behavior and Human Decision Processes*, 65(2), 148-158.

Zeelenberg, M., & Pieters, R. (2004). Consequences of regret aversion in real life: The case of the Dutch postcode lottery. *Organizational Behavior and Human Decision Processes*, 93(2), 155-168.

Zeelenberg, M., & Pieters, R. (2007). A theory of regret regulation 1.0. *Journal of Consumer Psychology*, 17(1), 3-18.

CHAPTER 7

General discussion

The primary research questions of this PhD thesis were: what is the 1) the short-term and 2) long-term effectiveness of different commitment lotteries in supporting lifestyle decisions that are in line with people's own goals? And 3) what are the contextual and psychological factors that help explain and optimize their effect and design? This PhD thesis further aimed to contribute to the science, policy and practice of prevention in a novel field setting. In this chapter, first, main findings are discussed. Second, scientific reflections are made. Third, the policy and practical implications are discussed. Fourth, future research directions are suggested. Finally, recommendations for science, policy and practice are given.

Main findings

In Chapter 2, participants reported their expected emotional response after missing out on a prize in one of twelve randomly presented incentive-scenarios, which varied in incentive type, incentive size and deadline distance. Participants primarily reported feeling disappointment, followed by regret. The likelihood and intensity of disappointment increased in the case of a traditional, fixed incentive. Regret was expected most when losing a lottery prize (vs. a fixed incentive) and intensified with prize size.

Emotion research has demonstrated that disappointment and regret differ in their level of agency, which has important behavioral implications (Frijda, 2007). Regret activates an action tendency to correct or prevent mistakes, because people who feel regret feel more responsible for a negative outcome. Disappointed people feel more powerless and have a stronger tendency to get away from the situation (Zeelenberg, Van Dijk, Manstead, & Van der Pligt, 2000). Disconfirmed expectations (i.e., disappointment) may result in attrition and eventually harm the success of interventions (Grave et al., 2005). In Chapter 2, it was reasoned that disappointment may have been widely reported because the scenarios described unsuccessful weight loss, which is not a decision but an undesired outcome that cannot be *directly* influenced by the individuals. This insight was the basis of an important decision for the design of commitment lotteries.

Chapter 3 described the design and protocol of the cluster randomized trial. Initially, it was planned to link the lotteries to successful weight loss. However, the aim of this endeavor was to use regret aversion to build a commitment device that would successfully support people's own goals. Based on scientific literature, it was reasoned that people most likely anticipate regret if they can directly link their decision to the negative outcome, and more unlikely when the path from decisions to negative outcomes is less volitional and more ambiguous. Because disappointment was reported so broadly in Chapter 2, it was decided to link the lotteries to a concrete 'wright or wrong' outcome that was under more direct volitional control than weight loss: gym attendance. Due to a progressive understanding of the emotional responses to the lotteries, the switch from supporting outcomes to supporting concrete decisions was made.

We aimed to examine if lottery-based deadlines that leverage regret aversion would help overweight individuals in attaining their goal of attending their gym twice per week. After determining the trial logistics and protocol, six corporate gyms (clusters) with a total of 163 overweight participants were randomized to one of three arms. In a parallel group, single-blinded, cluster randomized trial, three arms were followed for 52 weeks. The *short-term lottery arm* participated in weekly short-term lotteries for 13 weeks; the *long-term lottery arm* participated in identical short-term lotteries in combination with an additional long-term lottery after 26 weeks. Participants in the control arm also set the goal to attend their gym twice per week and were also actively monitored, but did not participate in the lotteries.

In Chapter 4, it was examined if the lotteries promoted physical activity among overweight adults up to 26 weeks and if the short-term lottery and long-term lottery arm yielded different effects. In the first 13 weeks, participants in both lottery arms attained their attendance goals significantly more often than participants in the control arm. Between weeks 13 and 26, we observed a decline in goal attainment in the lottery arms. Moreover, the highest goal attainment was found in the long-term lottery arm. Weekly short-term lotteries supported regular physical activity for 13 weeks and an additional long-term lottery after 26 weeks partly averted the decline in goal attainment after the 13 weekly lotteries.

After all the lotteries ended, participants were followed up until 52 weeks. Chapter 5 describes that after a successful 26-week intervention, goal attainment declined between weeks 27 and 52 in all arms and remained higher in the long-term lottery arm than in the control group. Goal attainment did not differ between the short-term lottery arm and control arm. Weight patterns over the course of the trial were also analyzed. Weight declined slightly in all arms in the first 13 weeks of the trial and remained stable from there on.

In Chapter 6, counterfactual feedback on delayed and uncertain gains was manipulated in three experiments to test if this would affect I) certainty equivalents and II) degrees of delay discounting. There was no compelling evidence of feedback attenuating delay discounting. However, as in the cluster randomized trial, counterfactual feedback affected decisions in the present with delayed consequences. In a choice between an immediate certain gain and a delayed uncertain gain, participants who were guaranteed to always learn the outcome of the delayed gamble valued the gamble higher. This effect was not significant when amounts were higher.

Scientific reflections

Self-control

The results presented in this PhD thesis indicate that psychological and behavioral economic insights can be of use in overcoming self-control troubles. Traditional models of behavior change typically focused on factors that could change individual trade-offs of costs and benefits and eventually behavioral intentions (Ajzen, 1991; Von Neumann & Morgenstern, 1944, 2007). However, a meta-analysis of experiments that strengthened intentions illustrates that a medium-to-large-sized change in intentions led to only a small-to-medium-sized change in behavior (Webb & Sheeran, 2006). The behavioral economic insight of dynamic inconsistency (choosing more deliberately for the future and more impulsively for the present) points to self-control problems and could help explain why people do not always act on their intentions; we weigh our preferences differently over time and we generally overweigh the present (Kirby & Herrnstein, 1995).

Exerting strong willpower is an intuitively appealing method to inhibit present-biased impulses and to bridge the gap between intention and behavior. Most psychological literature has indeed focused on ways to measure and increase the human strength of will (see: Marina & Michael, 2017). However, an expanding number of studies show that in the face of temptations, effortful restraint (deliberately overruling impulses) is weakly or unrelated to goal attainment (Fujita, 2011; Galla & Duckworth, 2015; Marina & Michael, 2017). Instead, people who show good self-control in daily life actually engage in *less* impulse control (Hofmann, Baumeister, Förster, & Vohs, 2012). Behavioral scientific literature increasingly demonstrates that effective self-control is more a matter of reducing temptation than of fighting temptation. People who act on their intentions are especially strong in strategically structuring their decision-environments in a way that reduces future temptations (Bryan, Karlan, & Nelson, 2010; Galla & Duckworth, 2015; Laibson, 2015; Steel, Brothen, & Wambach, 2001; Thaler & Benartzi, 2004).

Commitment lotteries were also designed for this purpose. They were aimed at helping people who had the intention of exercising on a regular basis and felt that they might not act on this. Without having to persuade people of the importance of exercise or significantly increasing the benefits (e.g., a traditional incentive), lottery deadlines served as a strategically structured decision-context for people to reduce the relative temptation of procrastination.

Building on psychological insights

In this attempt, commitment lotteries aimed to capitalize on multiple predictable deviations from the rational choice paradigm. First, we imposed nearby deadlines to draw the consequences of procrastination nearer. Voluntarily accepting a costly deadline limits the time until a goal can be attained, can reduce the relative appeal of procrastination and prioritize action (Ariely & Wertenbroch, 2002; Locke & Latham, 2006). Second, while we aimed to stimulate long-term

behavior change, we used week goals to fit individuals' myopia and to facilitate the human tendency of using temporal landmarks (e.g., Mondays) to relegate misfortune to the past and have a fresh start in the next week (Dai, Milkman, & Riis, 2014; O'Donoghue & Rabin, 1999). Third, a lottery is an effective and scalable tool to make missing the deadlines potentially costly for participants, because people generally do not rationally contemplate the odds of a lottery but tend to overestimate their chances of winning (Kahneman & Tversky, 1979). Fourth, we leveraged the human tendency to avoid regret by drawing the lottery prize from all participants and promising to inform non-eligible winners of their forgone prize (Zeelenberg & Pieters, 2004, 2007).

Regret aversion was targeted in commitment lotteries and studied more profoundly in this PhD thesis because this emotion has important behavioral implications. Psychological theories that capture the (evolutionary) function of emotions depict that emotions exist to prioritize our actions, mobilize our energy and serve our goals (Frijda, 2007; Haselton & Ketelaar, 2006). Because our feeling is for our doing, and emotions surface in predictable contexts (Zeelenberg, Nelissen, Breugelmans, & Pieters, 2008), emotions are well-suited to support individual goals.

The allure of using anticipated regret to support decision-making is that emotional processing requires considerably less cognitive capacity than the rational deliberation of the (intertemporal) pros and cons of a lifestyle decision (Frijda, 2007). The core value of contemporary behavioral science is the knowledge that the human ability to gather, compute and process information into an optimal (long-term) decision is limited in foreseeable situations. Emotions can help us overcome our systematic cognitive limitations because they automatically focus attention (Hanoch, 2002), assign value to decision-outcomes and energize us to obtain valued outcomes (Frijda, 1987; Zeelenberg et al., 2008).

Chapter 2 suggests a mixed pattern of emotions when a prize is forfeited. Regret indeed plays a role when participating in commitment lotteries and findings point to several psychological and contextual factors that influence regret. Haff et al. (2015) had also studied which demographics affected the effectiveness of lotteries with counterfactual feedback, but it remained unclear which features would emphasize a key motivational mechanism; regret aversion. Findings from Chapter 2 indicate several design features that can elicit and intensify regret, which could serve as the basis for further exploration of decision-contexts that influence emotions.

Findings from Chapter 6 further suggest that counterfactual feedback retains its influence if it is delayed. This may be of theoretical interest in regret research and in research of intertemporal decision-making. Because most lifestyle decisions have an intertemporal structure (e.g., exercise now for future health), counterfactual feedback might help attain our long-term goals. The experiments further indicate that if people have to reject a too-large, immediate and certain benefit for an uncertain and delayed outcome, merely feedback might not alter decisions. These insights may help design future commitment lotteries or other interventions

that aim to use counterfactual feedback to support safer (Koch, 2014) or healthier decisions (Brewer, DeFrank, & Gilkey, 2016).

Multiple disciplines and multiple angles

The development of commitment devices in the health domain is still in its early stages. Behavioral scientists are tweaking and comparing interventions to increase their effectiveness (Bryan et al., 2010; Rogers, Milkman, & Volpp, 2014). Commitment devices for health behavior are mostly tested in field experiments by multidisciplinary teams (e.g., medical doctors, psychologists, economists, epidemiologists; Loewenstein, Asch, Friedman, Melichar, & Volpp, 2012). Field experimentation by multidisciplinary teams appears to be a helpful approach to the tenacious challenge of health behavior change. Different disciplines raise different insights, which can be incorporated in field experiments to see what works.

This PhD thesis aimed to contribute to the understanding of health behavior by testing the effectiveness of multiple commitment lotteries in a cluster randomized trial and by attempting to grasp, from multiple angles, some of the behavioral processes that were targeted to improve the effectiveness of commitment lotteries. While the results from field experiments have the most direct link to policy and practice, it should not be overlooked that more fundamental insights in emotions and decision-processes are often needed for a solution-based applied behavioral science (Buunk & Van Vugt, 2013).

Practical reflections

Traditional methods

Lifestyle behavior is now a key determinant of our health (see Chapter 1). As a consequence, governments and other organizations aim to encourage people to adopt a healthy lifestyle. While subsidies, taxes or regulations are instinctively attractive to support lifestyle behaviors, these measures do not always have broad societal and political support, because the balance between paternalism and autonomy is complex and politically sensitive (Ten Have, 2014).

Consequently, non-regulatory methods to promote healthy lifestyle decisions have focused primarily on persuading or warning people via campaigns or other channels about the benefits or harms of certain behaviors (Marteau, Hollands, & Fletcher, 2012). For example, in their Action Plan for the Prevention and Control of Noncommunicable Diseases, the WHO recommends the provision of clear and independent information to enable citizens to choose wisely (WHO, 2011). Most likely, this strategy is implicitly or explicitly based on a rational view of human behavior: information persuades people to form a healthy intention, which they act upon (Prast, 2011). Indeed, in an economic publication by The Netherlands Scientific Council for Government Policy (WRR), unhealthy lifestyle decisions are described as an example of "information failures (...) through an inadequate understanding of the consequences of such behaviors" (Navarrete, De Visser, & Knottnerus, 2017. p. 20).

In line with this view, fear appealing warnings and images on cigarette packages are being used throughout the European Union to persuade smokers to quit. The knowledge of severe illness in the future should convince people to refrain from smoking (Kok, Ruiter, Van den Hoek, Schaalma, & De Vries, 2007). Likewise, in the food-domain, information about (un)healthy nutrients on points of purchase is widely used to stimulate smarter decisions (Grunert & Wills, 2007). After health-logos in Dutch supermarkets were removed from packaging because they were too confusing, the Dutch Minister of Health, Welfare and Sports ordered the development of a mobile application via which citizens could find independent nutritional information about their groceries. The question remains if these measures are sufficient.

Unfortunately, there is no indication, nor solid scientific basis to support the idea that fear-appealing warnings move smokers to abstinence (Kok, Peters, Kessels, Ten Hoor, & Ruiter, 2018; Kok et al., 2007). Furthermore, reviews of the literature suggest there is little to no evidence that nutritional information at points of purchase influences actual behavior (Grunert & Wills, 2007; Van't Riet, 2013). And while it is too early to tell whether the Dutch informational nutrition app actually results in healthier decisions, research into apps and wearable devices that monitor and inform people on their levels of physical activity demonstrates that these informational tools do not promote physical activity (Patel, Asch, & Volpp, 2015). Information is not a bad starting point, but is often insufficient in realizing behavior change.

Applied behavioral science

The integration of psychology and economics demonstrates that predictable deviations from rational behavior can hinder the translation from well-intended communication to behavior. In daily life, different goals compete for our attention and people simply lack the time and cognitive ability to collect, process and act on all relevant information (Simon, 1955). Fortunately, behavioral science provides a more realistic view of behavior that is being gradually recognized by organizations and governments. At an increasing rate, behavioral insights are assisting governments and organizations in adopting more effective regulatory solutions to help people realize their goals, without resorting to stringent laws or significant economic (dis)incentives (OECD, 2017a).

In the United Kingdom, the Behavioral Insights Team (BIT) celebrated its 8-year anniversary in 2018, and has since its establishment impressively improved economic, social and health policies with behavioral science. Following in its tracks, the previous president of the USA issued an executive order for Using Behavioral Science Insights to Better Serve the American People (Obama, 2015). Worldwide, BITs were formed to advise organizations and government agencies, including France, Germany, Denmark, the USA, Canada and the Netherlands (OECD, 2017a).

In the Netherlands, the WRR stressed in 2009 that policies should account for the factors that hinder the translation from knowledge to behavior (e.g., temptations and self-control). In their report "The human decision-maker", the WRR established that Dutch governments may be demanding too much rational self-reliance from their citizens, which interferes with what behavioral science teaches us about systematic bounded rationality (Tiemeijer, Thomas, & Prast, 2009). After a subsequent string of reports about the *potential* of applied behavioral science in the Netherlands, the Dutch coalition acknowledged explicitly in their 2018 Annual Budget that behavioral insights should actually be *used* to adopt a more realistic view of human behavior and improve the effectiveness of policies (Dijselbloem, 2017).

In the meantime, multiple Dutch organizations and governments were already experimenting or applying behavioral insights (including the Behavioral Insights Network Netherlands). A recent overview showed that at least 20 governmental agencies currently have ongoing behavioral insights programs (Feitsma, 2018). Accordingly, an inventory of applied behavioral science by public agencies in the Netherlands (BIN_NL, 2017) yielded multiple experiments by teams in the fields of taxes, finance, education, safety, and housing. Unfortunately, in the health domain, only one behavioral science experiment was reported; commitment lotteries.

Difficulties of lifestyle programs in practice

The benefits of regular physical activity are substantial; it lowers the risks of obesity, cardiovascular disease, type 2 diabetes, and breast and colon cancers (Lee et al., 2012). Yet, there are multiple possible reasons why the application of behavioral insights to the promotion of lifestyle behavior moves slower than in other domains (e.g., finance or consumer protection). First, if we assume that the allocation of resources reflects priority, treating disease is historically considered more important than preventing disease: preventive care expenditures were only 3.6% of the total Dutch health expenditure in 2016 (WHO, 2018). Second, while Dutch insurers and employers are offering prevention programs to their employees and clients respectively, it is unsure if they will ever see a proper return on investment, because people can switch jobs or insurers (i.e., the wrong-pocket problem; Heijink & Struijs, 2016). Third, individuals are free to make their own lifestyle decisions, even if these decisions harm their own health. It is politically complex and far from straightforward to determine to what degree insurers, doctors, employers or governments should be involved in our lifestyle decisions (Ten Have, 2014).

Fortunately, commitment lotteries might to some degree circumvent these issues. First, if we want to get as much health out of a euro as possible, because our resources are finite, behavioral economics can help. As Chapter 6 indicates, the mere promise of (delayed) counterfactual feedback increased the perceived value of money. Behavioral science has shown that our emotional, social, physical and temporal contexts affect our decisions, which can be used to increase the impact of money that is spent on prevention (Loewenstein et al., 2012). Furthermore,

in commitment lotteries, the majority of participants require no costs (84% did not receive a prize) and as mentioned in Chapter 4, by leveraging behavioral insights, we spent only €2.21 per participant per week (awarded prizes ÷ participants ÷ weeks) to stimulate physical activity for up to 52 weeks. That is only 0.6% of the Dutch minimum week-wage (€361.25).

Second, while health system payment reforms are gradually being implemented to overcome the wrong-pocket problem (Struijs, Hayen, & Van der Swaluw, 2018), we may not have to wait for this if participants would be willing to pay for commitment (ING, 2017). People already deposit their money into interest-free Christmas Clubs or impose penalties on their savings accounts to preempt the temptation of impulsive spending (Beshears et al., 2015; Thaler & Shefrin, 1981). Similarly, via a deposit contract, people put their own money at stake to attain their health goals (Halpern, Asch, & Volpp, 2012). If people would also be willing to pay for participation in a commitment lottery (as in a normal lottery), this might cover (some of the) costs, help the business case and bypass the wrong-pocket problem. Still, the willingness to pay for commitment could be low (Laibson, 2015) and the willingness to pay for commitment lotteries (and its effect on uptake) was not studied in this PhD thesis. This remains an interesting path for future inquiry.

Third, because opponents of lifestyle interventions fear a loss of autonomy, they argue that others (e.g., governments) should not limit their freedom of choice (Ten Have, 2014). Commitment devices and commitment lotteries avoid this issue by their voluntary nature. They are principally designed for autonomous individuals who are looking for ways to attain their own goals. Commitment lotteries did not oblige or fine (ir)regular exercise, nor did they bribe people towards the benefits of regular exercise; people were already informed and motivated. While an external party enforces people's commitments, this is typically the reason why forward-looking people impose a commitment device: they acknowledge that external enforcement may help them achieve their own goals. In line with this reasoning, the availability of commitment devices and the free choice to adopt or decline them can only be considered an expansion of freedom of choice and autonomy.

In the latest coalition agreement, the Dutch preventive policy goal is to primarily support evidence-based interventions. This brings the governmental responsibility of supporting field experimentation, innovation and pilot testing. Luckily, these are the primary tools of the behavioral scientist, which can be used to answer a number of open questions on commitment lotteries.

Future directions

Long term effectiveness

In Chapter 5, we find significant differences between the long-term lottery arm and the control arm at the 52-week follow-up. This shows a net effect of long-term commitment lotteries over the 52 weeks of the trial, and also suggests that regular gym attendance may have become a habit for some of the participants. Nonetheless, the quest for higher levels of maintenance remains ongoing. Habit formation is an effective way to strategically avoid temptations towards goal-attainment (Galla & Duckworth, 2015) and long-term commitment lotteries may have promoted habit formation to a degree. However, most conceptual and theoretical literature on commitment devices focusses on single tasks (e.g., Beshears et al., 2015; O'Donoghue & Rabin, 1999; Rogers et al., 2014; Thaler & Benartzi, 2004), while lifestyle is an ongoing process of repeated decisions.

Should we be concerned that, like medication, commitment devices generally lose most of their effect once people stop "taking them?", or should we follow the majority of the evidence showing that short-term measures generally yield short-term results and start focusing on structural, infinite opportunities for commitment? It would be valuable to know more conceptually about the effect of commitment devices on repeated decisions and which predictions existing theoretical models (e.g., Laibson, 2015) would make on this. This would help in the pursuit of long-term behavior change.

Motivation

A related open issue is the effect of commitment lotteries on motivation. In interventions that incorporate a financial component, there is often a concern that this may crowd out intrinsic motivation (Deci, Koestner, & Ryan, 1999). Crowding out would mean that participants in the intervention arms became primarily motivated by the lottery and eventually lost all interest in exercise. This would mean that, after the lotteries, lottery participants would attend their gym less often than the participants who had never participated in the lotteries (i.e., the control arm). Long-term analyses in Chapter 5 reveals no such pattern; although effects declined when the lotteries ended, there was no indication that levels of gym attendance became worse than they would have been if the lotteries had never been offered (i.e., lower than the control arm). This finding is in line with research that suggests that procrastination is not a motivational deficit (Steel et al., 2001). People with a tendency to postpone their efforts towards long-term goals form equal (or more) intentions, but just have more difficulty acting on them (Schouwenburg & Groenewoud, 2001). Nonetheless, the interaction of motivation and commitment devices remains unexplored. It would be valuable to know more about how commitment devices affect and interact with motivation and how we can design commitment lotteries that remain effective over repeated or perhaps infinite applications.

Implementation

Another important open matter is the implementation of commitment lotteries in practice. A common difficulty with prevention programs is that after being proven effective, they are underused in practice (Pryor & Volpp, 2018). One of the reasons is that physicians, insurers or employers, often have no suitable infrastructure (yet) to implement and manage evidence based interventions (Pryor & Volpp, 2018). A possible solution is the co-production and testing of interventions by science and practice. For the scientist, this can help identify everyday difficulties that may hinder practical success, and has the benefit that an intervention has potentially more external validity. For policy and practice, co-production and field testing can support the quest for evidence-based policies and relate what works (Halpern et al., 2018). Working together, science and practice can mutually benefit from each other's expertise and yield the best results for their population.

Uptake

In a recent survey by the New England Journal of Medicine Catalyst, health professionals stated that a lack of patient buy-in or engagement in preventive programs is the biggest obstacle to patient behavior change (Volpp & Seth Mohta, 2018). For every commitment device, there is a context-specific optimum between attractiveness and effectiveness (Halpern et al., 2012; Laibson, 2015). A very effective commitment device might be accepted by fewer people than a more attractive and moderately effective one, leading to a higher net effect of the moderately effective program for the targeted population. For example, Halpern and colleagues (2015) found that deposit contracts (participants lose their deposit if they fail) were considerably more effective than traditional rewards in assisting smoking cessation, but only a minority of participants accepted this deposit contract after random assignment. Ideally, the willingness to participate in different forms of commitment devices should be studied with the target population prior to the implementation of the program. This may help overcome the obstacles of engagement and uptake.

Another route to a broader use of commitment devices might be private paternalism. Laibson (2018) introduced the term private paternalism as policies by private institutions that "advance an individual's interests by restricting his or her freedom" (p2.). In this context, commitment is part of a larger set of measures that help people help themselves. Employers or gyms do not necessarily have to promote their commitment devices, but only their good results (e.g., high gym attendance rates). An example by Laibson is a good university that showcases its successful alumni, but not necessarily that these alumni were coaxed towards their academic achievements with strict deadlines.

A benefit of private paternalism is that forward looking individuals may elect it because it helps them overcome their self-control difficulties, and naïve people (who overestimate their future self-control) may choose it because they value the genuine narratives of good results. With this

reasoning, an ideal situation for people with self-control problems would be for commitment lotteries to be a standard part of a gym membership or of a firm's human resources policy. In that case, uptake is voluntary at lower frequencies (i.e., the choice for a gym or employer), but enforcing at high frequencies (i.e., weekly lotteries; Laibson, 2018). Empirical research should reveal whether this concept fits the goals of organizations and employees.

Heterogeneity

Similar to worldwide observations, Dutch citizens with a relatively lower education and lower income generally engage in more health damaging lifestyle behaviors and age in a less healthy way (RIVM, 2018). Besides, many preventive measures in the past reached higher income groups better than lower income groups (OECD, 2017b). Ideally, interpersonal differences should be taken into account in the designing stages of future commitment lotteries. Similar to a meta-analysis of multiple lotteries by Haff et al. (2015), we identified no differences in lottery effectiveness between income or education categories, which may indicate that commitment lotteries are widely applicable. Still, it is not unlikely that different programs are attractive and effective for different people (Halpern et al., 2015). The heterogeneity in reach of behavioral interventions is an important path for future research. More knowledge about the demographic characteristics that contribute to the effectiveness and uptake of commitment lotteries would be helpful.

Recommendations

For (teams of) researchers, policymakers, health professionals, insurers or employers, several recommendations can be made. First, in the effort to assist health behavior change, be aware that information although a good starting point, is rarely sufficient. Second, realize that unhealthy behaviors do not necessarily point to a lack of information, motivation or incentive, but that lapses of self-control are omnipresent.

Third, given the existence of self-control problems, people's interests can be served if they can restrict their freedom of choice. Empirical observations show that autonomous people value and also voluntarily impose choice restrictions (e.g., deadlines). But where can people go if they are looking for an effective commitment device for their lifestyle goals? There seems to be a lack of governmental and organizational tools for people to voluntarily restrict their lifestyle decisions (e.g., with respect to food, exercise, or purchasing alcohol and cigarettes). In order to meet the world's ambitious goals to reduce the burden of noncommunicable diseases, it is recommended that (teams of) researchers, policy makers, health professionals, insurers or employers start offering (and remain testing) commitment devices.

Fourth, when designing interventions to help people help themselves, think of using emotions to bypass cognitive limitations, but be aware that events can evoke multiple emotions that may not always serve people's goals. Fifth, when designing a commitment lottery, focus on features 1) that make lotteries attractive for people to say yes to 2) that yield success over repeated and possibly infinite applications and 3) that contribute to its implementation in practice.

REFERENCES

WHO, World health Organization. (2011). *Action plan for implementation of the European Strategy for the Prevention and Control of Noncommunicable Diseases 2012–2016*. Copenhagen: WHO Regional Office for Europe.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211.

Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: self-control by precommitment. *Psychological Science, 13*(3), 219-224.

Beshears, J., Choi, J. J., Harris, C., Laibson, D., Madrian, B. C., & Sakong, J. (2015). Self control and commitment: can decreasing the liquidity of a savings account increase deposits? *National Bureau of Economic Research, w21474*, 1–53.

BIN_NL, Behavioural Insights Netwerk Nederland. (2017). *Rijk aan gedragsinzichten: editie 2017*. Den Haag: Ministerie van Economische Zaken en Klimaat.

Brewer, N. T., DeFrank, J. T., & Gilkey, M. B. (2016). Anticipated regret and health behavior: A meta-analysis. *Health Psychology, 35*(11), 1264-1275.

Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment devices. *Annual Review of Economics, 2*(1), 671-698.

Buunk, A. P., & Van Vugt, M. (2013). *Applying social psychology: From problems to solutions*. California: Sage Publications.

Dai, H., Milkman, K. L., & Riis, J. (2014). The fresh start effect: Temporal landmarks motivate aspirational behavior. *Management Science, 60*(10), 2563-2582.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627.

Feitsma, J. (2018). Brokering behaviour change: the work of behavioural insights experts in government. *Policy & Politics (Online)*.

Frijda, N. H. (1987). Emotion, cognitive structure, and action tendency. *Cognition and Emotion, 1*(2), 115-143.

Frijda, N. H. (2007). *The laws of emotion*. Mahwah: Lawrence Erlbaum Associates Publishers.

Fujita, K. (2011). On conceptualizing self-control as more than the effortful inhibition of impulses. *Personality and Social Psychology Review, 15*(4), 352-366.

Galla, B. M., & Duckworth, A. L. (2015). More than resisting temptation: Beneficial habits mediate the relationship between self-control and positive life outcomes. *Journal of Personality and Social Psychology, 109*(3), 508-525.

Grave, R., Calugi, S., Molinari, E., Petroni, M. L., Bondi, M., Compare, A., & Marchesini, G. (2005). Weight loss expectations in obese patients and treatment attrition: an observational multicenter study. *Obesity Research, 13*(11), 1961-1969.

Grunert, K. G., & Wills, J. M. (2007). A review of European research on consumer response to nutrition information on food labels. *Journal of Public Health, 15*(5), 385-399.

Haff, N., Patel, M. S., Lim, R., Zhu, J., Troxel, A. B., Asch, D. A., & Volpp, K. G. (2015). The role of behavioral economic incentive design and demographic characteristics in financial incentive-based approaches to changing health behaviors: a meta-analysis. *American Journal of Health Promotion*, *29*(5), 314-323.

Halpern, S. D., Asch, D. A., & Volpp, K. G. (2012). Commitment contracts as a way to health. *BMJ: British Medical Journal*, *344*(1), e522-e522.

Halpern, S. D., French, B., Small, D. S., Saulsgiver, K., Harhay, M. O., Audrain-McGovern, J., . . . Volpp, K. G. (2015). Randomized trial of four financial-incentive programs for smoking cessation. *New England Journal of Medicine*, *372*(22), 2108-2117.

Halpern, S. D., Harhay, M. O., Saulsgiver, K., Brophy, C., Troxel, A. B., & Volpp, K. G. A pragmatic trial of e-cigarettes, incentives, and drugs for smoking cessation. *New England Journal of Medicine*, *378*, 2302-2310.

Hanoch, Y. (2002). "Neither an angel nor an ant": Emotion as an aid to bounded rationality. *Journal of Economic Psychology*, *23*(1), 1-25.

Haselton, M. G., & Ketelaar, T. (2006). Irrational emotions or emotional wisdom? The evolutionary psychology of affect and social behavior. In J. P. Forgas (Ed.), *Affect in social thinking and behavior* (pp. 21-39). New York: Psychology Press.

Heijink, R., & Struijs, J. (2016). *Preventie in het zorgstelsel: wat kunnen we leren van het buitenland? Een eerste inventarisatie op basis van interviews met buitenlandse experts en literatuur.* Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu.

Hofmann, W., Baumeister, R. F., Förster, G., & Vohs, K. D. (2012). Everyday temptations: An experience sampling study of desire, conflict, and self-control. *Journal of Personality and Social Psychology*, *102*(6), 1318-1335.

ING. (2017). *Goede voornemens 2018*. Retrieved from www.ing.nl/media/A3%20Goede%20voornemens_tcm162-137795.pdf

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47*(2) ,263-291.

Kirby, K. N., & Herrnstein, R. J. (1995). Preference reversals due to myopic discounting of delayed reward. *Psychological Science*, *6*(2), 83-89.

Koch, E. J. (2014). How does anticipated regret influence health and safety decisions? A literature review. *Basic and Applied Social Psychology*, *36*(5), 397-412.

Kok, G., Peters, G.-J. Y., Kessels, L. T. E., ten Hoor, G. A., & Ruiter, R. A. C. (2018). Ignoring theory and misinterpreting evidence: the false belief in fear appeals. *Health Psychology Review*, *12*(2), 111-125.

Kok, G., Ruiter, R., van den Hoek, K., Schaalma, H. P., & de Vries, N. K. (2007). Waarschuwingsplaatjes op tabaksverpakking zijn niet effectief. *TSG*, *85*(3), 133-135.

Laibson, D. (2015). Why don't present-biased agents make commitments? *American Economic Review Paper and Proceedings*, *105*(5), 267-272.

Laibson, D. (2018). Private Paternalism, the Commitment Puzzle, and Model-Free Equilibrium (Richard T. Ely Lecture). *American Economic Review Papers and Proceedings Forthcoming*.

Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet, 380* (9838), 219-229.

Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science, 15*(5), 265-268.

Loewenstein, G., Asch, D. A., Friedman, J. Y., Melichar, L. A., & Volpp, K. G. (2012). Can behavioural economics make us healthier? *BMJ: British Medical Journal (Online), 344, e3482*.

Van den Berg, M. & Schoemaker, C.G. (2010). *Effecten van preventie. Deelrapport van de VTV 2010: Van gezond naar beter.* Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu.

Marina, M., & Michael, I. (2017). What's so great about self-control? Examining the importance of effortful self-control and temptation in predicting real-life depletion and goal attainment. *Social Psychological and Personality Science, 8*(6), 603-611.

Marteau, T. M., Hollands, G. J., & Fletcher, P. C. (2012). Changing human behavior to prevent disease: the importance of targeting automatic processes. *Science, 337*(6101), 1492-1495.

Navarrete, L. I. F., De Visser, M., & Knottnerus, J. A. (2017). *Societal and economic value of prevention in the Netherlands. Attaining happier, healthier and more productive lives.* Den Haag: The Netherlands Scientific Council for Government Policy.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review, 89*(1), 103-124.

Obama, B. H. (2015). *Executive Order - Using Behavioral Science Insights to Better Serve the American People*. Washington: The White House - Office of the Press Secretary.

OECD. (2017a). *Behavioural Insights and Public Policy: Lessons from Around the World*: Paris: OECD Publishing.

OECD and European Observatory on Health Systems and Policies (2017b). *State of Health in the EU. Netherlands: Country Health Profile 2017*. Paris: OECD Publishing.

Patel, M. S., Asch, D. A., & Volpp, K. G. (2015). Wearable devices as facilitators, not drivers, of health behavior change. *JAMA, 313*(5), 459-460.

Prast, H. M. (2011). A behavioral economics approach to public health. *European Journal of Consumer Law, 4*, 723 - 747.

Pryor, K., & Volpp, K. (2018). Deployment of Preventive Interventions—Time for a Paradigm Shift. *New England Journal of Medicine, 378*(19), 1761-1763.

Rogers, T., Milkman, K. L., & Volpp, K. G. (2014). Commitment devices: Using initiatives to change behavior. *JAMA, 311*(20), 2065-2066.

Schouwenburg, H. C., & Groenewoud, J. (2001). Study motivation under social temptation; effects of trait procrastination. *Personality and Individual Differences, 30*(2), 229-240.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69*(1), 99-118.

Dijselbloem, J. R.V. A. (2017). *Miljoenennota 2018*. Den Haag: Tweede Kamer der Staten-Generaal.

Steel, P., Brothen, T., & Wambach, C. (2001). Procrastination and personality, performance, and mood. *Personality and Individual Differences, 30*(1), 95-106.

Struijs, J., Hayen, A., & Van der Swaluw, K. (2018). When Designing Bundled Payments, Don't Ignore The Lessons Of Behavioral Economics [blog]. Retrieved from www.healthaffairs.org/do/10.1377/hblog20180420.640240/full/

Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy*, 112(S1), S164-S187.

Ten Have, M. (2014). *Leefstijlbeïnvloeding: tussen betuttelen en verwaarlozen. Signalering Ethiek en Gezondheid 2014/1.* Den Haag: Centrum voor Ethiek en Gezondheid.

Thaler, R. H., & Shefrin, H. M. (1981). An economic theory of self-control. *The Journal of Political Economy*, 89 (2), 392-406.

Tiemeijer, W. L., Thomas, C. A., & Prast, H. M. (2009). *De menselijke beslisser - over de psychologie van keuze en gedrag*. Den Haag, Amsterdam: Wetenschappelijke Raad voor het Regeringsbeleid.

Van't Riet, J. (2013). Sales effects of product health information at points of purchase: a systematic review. *Public Health Nutrition*, 16(3), 418-429.

Volpp, K. G., & Seth Mohta, N. (2018). *Patient Engagement Survey. What Creates Behavior Change May Not Sustain It.* Walthalm: Massachusetts Medical Society.

Von Neumann, J., & Morgenstern, O. (2007). *Theory of games and economic behavior (60th Anniversary Commemorative Edition)*. New Jersey: Princeton university press.

Webb, T. L., & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, 132(2), 249-268.

WHO, World Health Organization (2018). *Organization and financing of public health services in Europe: country reports (2018)*. Geneva: WHO.

Zeelenberg, M., Nelissen, R. M., Breugelmans, S. M., & Pieters, R. (2008). On emotion specificity in decision making: Why feeling is for doing. *Judgment and Decision Making*, 3(1), 18-27.

Zeelenberg, M., & Pieters, R. (2004). Consequences of regret aversion in real life: The case of the Dutch postcode lottery. *Organizational Behavior and Human Decision Processes*, 93(2), 155-168.

Zeelenberg, M., & Pieters, R. (2007). A theory of regret regulation 1.0. *Journal of Consumer Psychology*, 17(1), 3-18.

Zeelenberg, M., van Dijk, W. W., Manstead, A. S. R., & Van der Pligt, J. (2000). On bad decisions and disconfirmed expectancies: The psychology of regret and disappointment. *Cognition and Emotion*, 14(4), 521-541.

Summary

Behavior is a key determinant of health. Therefore, disease prevention through improved lifestyle behavior is receiving increasing attention from policymakers, medical professionals and science. Likewise, the fields of health economics, psychology and behavioral economics progressively understand the determinants of health related behavior. Knowledge of the latter is essential for public health policy and practice. Operating at the crossroads of health economics, psychology and behavioral economics, this PhD thesis focuses on supporting individual lifestyle decisions.

In the Netherlands, 18.5% of healthy life years lost is linked to behavior: after smoking, unhealthy nutrition and physical inactivity are the two leading contributors. Half of the Dutch population is overweight (Body Mass Index ≥ 25), 53% does not meet recommended levels of physical activity and about 43% exercises less than once per week.
Many people aim to change their lifestyle, but have trouble acting on their intentions. This PhD thesis focusses on people who are aware of the benefits of a healthy lifestyle, who want to realize their personal behavior change, but also feel that they may not act on this in the future.

It was studied whether commitment lotteries would support people in their goal to exercise on a regular basis. In a commitment lottery, participants set a behavioral health goal, to be achieved at a prespecified deadline. On the deadline, a prize is drawn out of all participants and announced to all. Importantly, the winners are only eligible for their prize if they attained their personal goal. As a consequence, non-eligible winners are informed about their forgone earnings. This counterfactual feedback is designed to provoke anticipated regret and emphasize the lottery deadlines.

The primary research questions of this PhD thesis were: what is the 1) short-term and 2) long-term effectiveness of different commitment lotteries in supporting lifestyle decisions that are in line with people's own goals? And 3) what are the contextual and psychological factors that help explain and optimize their effect and design? This PhD thesis further aimed to contribute to the science, policy and practice of prevention in a novel field setting.

In **Chapter 2**, participants reported their expected emotional response after missing out on a prize. It was explored which emotions were actually expected and which incentive-characteristics influence their *likelihood* and *intensity.* Participants reported their expected emotional response after missing out on a prize in one of 12 randomly presented incentive-scenarios, which varied in incentive type, incentive size and deadline distance. Participants primarily reported feeling disappointment, followed by regret. Regret was expected most when losing a lottery prize (vs. a fixed incentive) and intensified with prize size. Multiple features of the participant and the lottery incentive increased the occurrence and intensity of regret. These findings could be used to design the field experiment in this thesis.

Chapter 3 described the design and protocol of the field experiment. The aim of this experiment was to test whether a lottery-based commitment device could promote regular gym attendance. The winners of the lottery always get feedback on the outcome but can only claim their prize if they attended their gyms on a regular basis (gym attendance ≥ 2 per week). This chapter presents the design and baseline characteristics of a three-arm trial. After determining the trial logistics and protocol, six corporate gyms (clusters) with a total of 163 overweight participants were randomized to one of the three arms. All participants had the goal to attend their gym at least twice per week. In a parallel group, single-blinded, cluster randomized trial, the three arms were followed for 52 weeks. The short-term lottery arm participated in weekly short-term lotteries for 13 weeks (each worth €100); the long-term lottery arm participated in identical short-term lotteries in combination with an additional long-term lottery after 26 weeks (a luxurious family-vacation). After 26 weeks, all participants in the long-term lottery arm could win the vacation, but the winner could only keep the prize if he or she had attained the week-goal in at least 9 weeks (70%) between weeks 14 and 26. Participants in the control arm also set the goal to attend their gym twice per week and were also actively monitored, but did not participate in the lotteries.

In **Chapter 4**, it was examined if commitment lotteries promoted physical activity among overweight adults up to 26 weeks and if the short-term lottery and long-term lottery arm yielded different effects. We compared 1) weekly short-term lotteries for 13 weeks; 2) the same short-term lotteries in combination with an additional long-term lottery after 26 weeks; and 3) a control arm without lotteries. After 13 weeks, participants in the lottery arms attained their attendance-goals more often than participants in the control arm. After 26 weeks, we observe a decline in goal attainment in the short-term lottery arm and the highest goal attainment in the long-term lottery arm. Weekly short-term lotteries supported regular physical activity for 13 weeks and an additional long-term lottery after 26 weeks partly averted the decline in goal attainment after the 13 weekly lotteries.

After all the lotteries ended, participants were followed up until 52 weeks. In **Chapter 5**, maintenance of goal attainment at 52-week follow-up and the development of weight over time were analyzed. We compared weight and goal attainment (gym attendance ≥ 2 per week) between the three arms. After a successful 26-week intervention, goal attainment declined between weeks 27 and 52 in all arms and remained higher in the long-term lottery arm than in the control group. Goal attainment did not differ between the short-term lottery arm and control arm. Weight declined slightly in all arms in the first 13 weeks of the trial and remained stable from there on. Commitment lotteries can support regular gym attendance up to 52 weeks and more research is needed to achieve higher levels of maintenance and weight loss.

If decision-makers expect to learn the outcome of a non-chosen decision-alternative (i.e., expect to receive counterfactual feedback), this impacts their decisions. It remained

unclear to what extent such feedback retains its influence if feedback is delayed and whether feedback influences delay discounting. In **Chapter 6**, counterfactual feedback on delayed and uncertain gains was manipulated in three experiments to test if this would affect I) certainty equivalents and II) degrees of delay discounting. There was no compelling evidence of feedback attenuating delay discounting. However, as in the cluster randomized trial, counterfactual feedback affected decisions in the present with delayed consequences. In a choice between an immediate certain gain and a delayed uncertain gain, participants who were guaranteed to always learn the outcome of the delayed gamble valued the gamble higher. This effect was not significant when amounts were higher. Counterfactual feedback may be effective in supporting present decisions with delayed uncertain consequences.

DISCUSSION

The results presented in this PhD thesis indicate that psychological and behavioral economic insights can be of use in overcoming self-control troubles. Without having to persuade people of the importance of exercise or significantly increasing the benefits (e.g., a traditional incentive), lottery deadlines served as a commitment device for people to attain their own goals.

The allure of using anticipated regret to support decision-making is that emotional processing requires considerably less cognitive capacity than the rational deliberation of the (intertemporal) pros and cons of a lifestyle decision. The core value of contemporary behavioral science is the knowledge that the human ability to gather, compute and process information into an optimal (long-term) decision is limited in foreseeable situations. Emotions can help us overcome our systematic cognitive limitations because they automatically focus attention, assign value to decision-outcomes and energize us to obtain valued outcomes.

Most attempts to promote healthy lifestyle decisions have focused primarily on persuading or warning people via campaigns or other channels about the benefits or harms of certain behaviors. Although information is not a bad starting point, it is often insufficient in realizing behavior change.

The integration of psychology and economics demonstrates that predictable deviations from rational behavior can hinder the translation from well-intended communication to behavior. In daily life, different goals compete for our attention and people simply lack the time and cognitive ability to collect, process and act on all relevant information. Fortunately, behavioral science provides a more realistic view of behavior that is being gradually recognized by organizations and governments. Yet, the application of behavioral insights to the promotion of lifestyle behavior moves slower than in other domains (e.g., finance or consumer protection).

Fortunately, commitment lotteries might to some degree circumvent some of the most common barriers.

For (teams of) researchers, policymakers, health professionals, insurers or employers, several recommendations can be made. First, in the effort to assist health behavior change, be aware that information although a good starting point, is rarely sufficient. Second, realize that unhealthy behaviors do not necessarily point to a lack of information, motivation or incentive, but that lapses of self-control are omnipresent.

Third, given the existence of self-control problems, people's interests can be served if they can restrict their freedom of choice. Empirical observations show that autonomous people value and also voluntarily impose choice restrictions (e.g., deadlines).

In order to meet the world's ambitious goals to reduce the burden of noncommunicable diseases, it is recommended that (teams of) researchers, policy makers, health professionals, insurers or employers start offering (and remain testing) commitment devices.

Fourth, when designing interventions to help people help themselves, think of using emotions to bypass cognitive limitations, but be aware that events can evoke multiple emotions that may not always serve people's goals. Fifth, when designing a commitment lottery, focus on features 1) that make lotteries attractive for people to say yes to 2) that yield success over repeated and possibly infinite applications and 3) that contribute to its implementation in practice.

Nederlandse samenvatting

Onze gezondheid is voor een belangrijk deel afhankelijk van ons gedrag. Daarom krijgt ziektepreventie via leefstijlverandering steeds meer aandacht van beleidsmakers, medisch professionals en de wetenschap. Tegelijkertijd begrijpen de wetenschappelijke velden gezondheidseconomie, psychologie en gedragseconomie steeds meer van de determinanten van gezondheidsgedrag. Deze kennis is essentieel voor de praktijk en het beleid van publieke gezondheid. Op het snijvlak tussen gezondheidseconomie, psychologie en gedragseconomie richt dit proefschrift zich op het ondersteunen van leefstijlkeuzes.

In Nederland is 18.5% van de ziektelast gerelateerd aan gedrag: na roken zijn ongezond eten en te weinig bewegen de twee grootste veroorzakers. De helft van Nederland heeft overgewicht (Body Mass Index ≥ 25), 53% voldoet niet aan de beweegrichtlijnen en 43% sport minder dan een keer per week.

Veel mensen willen hun leefstijl verbeteren, maar vinden het lastig om zich naar hun intenties te gedragen. In dit proefschrift ligt de nadruk op mensen die zich bewust zijn van de voordelen van een gezonde leefstijl, die ook het doel hebben hun persoonlijke gedragsverandering te verwerkelijken, maar zich ook realiseren dat ze zich hier in de toekomst misschien niet naar zullen gedragen.

Voor dit proefschrift is onderzocht of een beweegloterij mensen kan ondersteunen bij hun doel om regelmatig te sporten. In de beweegloterij stellen mensen een gedragsdoel, dat gehaald dient te zijn op een vooraf gestelde deadline. Op de deadline wordt er een prijs verloot onder alle deelnemers en wordt de winnaar aan iedereen bekend gemaakt. Echter, de winnaars krijgen hun prijs alleen als ze ook hun doel hebben gehaald. Hieruit volgt dat winnaars die hun doel niet hebben gehaald, onvermijdelijk te horen krijgen wat ze zijn misgelopen. Deze counterfactual feedback (terugkoppeling over verdane kansen) is bedoeld om deelnemers te laten anticiperen op spijt, en zo de deadlines te benadrukken.

De primaire onderzoeksvragen van dit proefschrift waren: wat is de 1) kortetermijn- en 2) langetermijneffectiviteit van verschillende beweegloterijen in het ondersteunen van leefstijlkeuzes die voortkomen uit mensen hun eigen doelen? En 3) wat zijn de contextuele en psychologische factoren die het effect en het ontwerp van de beweegloterij helpen verklaren en optimaliseren? Dit proefschrift was daarnaast bedoeld om bij te dragen aan de wetenschap, de praktijk en het beleid van preventie in een nieuwe praktijkomgeving.

In **Hoofdstuk 2** rapporteerden onderzoeksdeelnemers hun verwachte emotionele reactie op het mislopen van een prijs. Er werd onderzocht welke emoties worden verwacht, en welke factoren invloed hebben op hun waarschijnlijkheid en intensiteit. Nadat deelnemers 1 van 12 willekeurig gepresenteerde scenario's lazen waarin ze een prijs misliepen, werd hun gevraagd welke emotie ze zouden voelen en in welke mate. De scenario's varieerden in prijstype, prijsgrootte en de afstand tot de deadline. De deelnemers rapporteerden vooral teleurstelling, gevolgd

door spijt. Spijt werd vaker genoemd bij het mislopen van een loterij (versus een gegarandeerde prijs) en werd intenser met het toenemen van de prijsgrootte. Er zijn meerdere karakteristieken van de persoon en het prijstype, die bijdragen aan de waarschijnlijkheid en de intensiteit van spijt. Deze bevingen konden worden gebruikt om het veldexperiment in dit proefschrift te ontwerpen.

In **Hoofdstuk 3** worden het ontwerp en het protocol van het veldexperiment beschreven. Het doel van dit experiment was onderzoeken of een zelfbindingsmechanisme in de vorm van een beweegloterij regelmatig sporten zou kunnen bevorderen. De winnaars krijgen altijd te horen dat ze gewonnen hebben, maar krijgen hun prijs alleen als ze ook hun beweegdoel hebben gehaald (twee of meer sportschoolbezoeken per week). Nadat de logistiek en het protocol bepaald waren, zijn zes sportscholen (clusters) met totaal 163 deelnemers met overgewicht willekeurig toegewezen aan een van drie onderzoeksarmen. Alle deelnemers stelden het doel om minimaal twee keer per week te sporten.

In een cluster gerandomiseerde trial werden de drie armen 52 weken gevolgd. De *kortetermijngroep* deed 13 weken lang wekelijks mee aan een loterij (ter waarde van €100). Alle deelnemers konden iedere week opnieuw winnen, maar kregen hun prijs alleen als ze in de betreffende week het doel gehaald hadden (twee of meer sportschoolbezoeken per week). De *langetermijngroep* deed in de eerste 13 weken mee aan 13 identieke loterijen en deed daarna mee aan nóg een loterij, die na 26 weken plaatsvond. Hier werd een luxe gezinsvakantie verloot. Wederom werd de prijs verloot onder alle deelnemers en altijd bekend gemaakt. De winnaar mocht de prijs echter alleen houden als hij of zij in minstens 9 weken het weekdoel had gehaald (70% tussen weken 14 en 26).

De *controlegroep* werd ook actief gemonitord. In de controlegroep stelden deelnemers ook het doel om minimaal twee keer per week naar de sportschool te gaan, maar vonden er geen loterijen plaats.

In **Hoofdstuk 4** wordt onderzocht of de beweegloterij effectief was tot 26 weken, en of de twee verschillende loterijen andere resultaten opbrachten. Er wordt een vergelijking gemaakt tussen 1) 13 wekelijkse loterijen 2) 13 wekelijkse loterijen die werden aangevuld met een loterij na 26 weken 3) de controlegroep. In de eerste 13 weken haalden de *loterijgroepen* hun doel significant vaker dan de *controlegroep*. Tussen weken 14 en 26 vond er een afname plaats in het aantal succesvolle weken. De *kortetermijngroep* verschilde in deze periode niet langer van de controlegroep in de kans om het doel te halen. De *langetermijngroep* presteerde wel significant beter dan de controlegroep. Er wordt geconcludeerd dat dertien wekelijkse loterijen het regelmatig sporten 13 weken lang ondersteunde, en daarna niet meer. Een additionele loterij na 26 weken ging de daling in sportschoolbezoeken grotendeels tegen.

Nadat alle loterijen afgelopen waren, werden alle deelnemers gevolgd tot aan 52 weken. In **Hoofdstuk 5** wordt geanalyseerd in hoeverre deelnemers het sporten tussen weken 26 en 52 volhielden en of hun gewicht over de 52 weken heen afgenomen was. Er wordt een

vergelijking gemaakt in gewicht en doelsucces (twee of meer keer per week sporten) tussen de drie onderzoeksgroepen. Na een succesvolle 26-weekse interventie daalt het aantal sportschoolbezoeken in alle groepen. In de *langetermijngroep* halveert de kans op het halen van het doel, maar die kans is nog steeds significant hoger dan in de *controlegroep*. De *kortetermijngroep* verschilt niet van de controlegroep. Het gewicht van de deelnemers daalt een klein beetje in alle groepen in de eerste 13 weken en blijft daarna stabiel. De beweegloterij kan regelmatig sporten tot 52 weken ondersteunen en er is meer onderzoek nodig om hogere volhoudpercentages en gewichtsverlies te bereiken.

Als mensen bij het maken van een keuze verwachten dat ze altijd de uitkomst van de niet-gekozen keuze-optie te horen krijgen, heeft dit invloed op hun keuzes. Het was echter onduidelijk wat het effect van dergelijke feedback zou zijn als er tijd zit tussen de keuze en de feedback. In **Hoofdstuk 6** werd in drie experimenten feedback beloofd op onzekere opbrengsten in de toekomst. Er werd onderzocht of deze feedback effect had op de 1) waardering en de 2) tijdsverdiscontering van vertraagde en onzekere opbrengsten.

Feedback had geen invloed op de verdiscontering van vertraagde opbrengsten. Echter, net als in het veldexperiment beïnvloede de garantie van feedback de keuzes met toekomstige gevolgen. Bij het kiezen tussen een onmiddellijke en zekere opbrengst versus een gok in de toekomst, waarderen mensen de toekomstige gok hoger als ze weten dat ze altijd (ongeacht hun keuze) de uitkomst van de gok te horen zullen krijgen. Dit effect werd niet gevonden als de bedragen hoger waren. Counterfactual feedback is mogelijk effectief in het ondersteunen van huidige keuzes met onzekere en vertraagde gevolgen.

DISCUSSIE

De resultaten in dit proefschrift wijzen erop dat psychologische en gedragseconomische inzichten kunnen helpen bij het omgaan met zelfcontroleproblemen. Zonder dat mensen extra overtuigd hoefden te worden over de voordelen van sport en zonder de opbrengsten van het sporten substantieel te vergroten (bijv. een traditionele financiële prikkel), hielpen de loterijdeadlines als zelfbindingsmechanisme mensen hun eigen doelen te halen.

De aantrekkingskracht van het gebruik van geanticipeerde spijt om keuzes te ondersteunen is dat de verwerking van emoties aanzienlijk minder cognitieve capaciteit kost dan de rationele calculatie van de (intertemporele) voor- en nadelen van een keuze. De kernwaarde van hedendaagse gedragswetenschap is de kennis dat het menselijk vermogen om informatie te verzamelen, te ordenen en te verwerken tot een optimale langetermijnbeslissing beperkt is op voorspelbare momenten. Emoties kunnen ons helpen onze systematische cognitieve beperkingen tegen te gaan, omdat emoties automatisch onze aandacht sturen, waarde geven aan keuze-opties en ons activeren om gewenste doelen te bereiken.

De meeste pogingen om een gezonde leefstijl te bevorderen richten zich voornamelijk op het gebruik van campagnes om mensen te overtuigen of te waarschuwen over de kosten en opbrengsten van bepaald gedrag. Hoewel informatie verstrekken geen slecht startpunt is, is het vaak ontoereikend om gedrag te veranderen.

De integratie van psychologie en economie laat zien dat voorspelbare afwijkingen van rationeel gedrag de transformatie van goedbedoelde communicatie naar gedrag kunnen belemmeren. In het dagelijks leven concurreren meerdere doelen om onze aandacht en mensen hebben simpelweg niet de tijd en cognitieve capaciteit om alle relevante informatie te verzamelen, te verwerken en ernaar te handelen.

Gelukkig biedt de gedragswetenschap een realistischer mensbeeld dat geleidelijk aan wordt erkend door organisaties en overheden. Desondanks worden gedragswetenschappelijke inzichten minder snel toegepast bij leefstijlverandering dan in andere domeinen (bijvoorbeeld financieel toezicht of consumentenbescherming). De beweegloterij kan mogelijk enkele van de remmende factoren omzeilen.

Voor (teams van) onderzoekers, beleidsmakers, gezondheidsprofessionals, verzekeraars of werkgevers zijn er enkele aanbevelingen. Ten eerste, in een poging om mensen te helpen hun leefstijl te verbeteren, is het goed ervan bewust te zijn dat informatie en bewustzijn op zichzelf zelden gedragsverandering teweeg brengen. Als tweede, realiseer dat ongezonde keuzes niet noodzakelijk wijzen op een gebrek aan kennis, motivatie of prikkel, maar dat zelfcontroleproblemen erg vaak voorkomen.

Als derde, gegeven het bestaan van zelfcontroleproblemen kunnen mensen baat hebben bij vrijwillige inperking van hun eigen keuzevrijheid. Empirie laat zien dat volledig autonome mensen keuzebeperkingen waarderen en zichzelf opleggen om hun eigen doelen te halen.
Om de wereldwijde ambities van het terugdringen van de chronische ziektelast te realiseren, is het sterk aanbevolen dat (teams van) onderzoekers, beleidsmakers, gezondheidsprofessionals, verzekeraars of werkgevers zelfbindingsmechanismen (commitiment devices) verder onderzoeken en gaan aanbieden.

Als vierde, bij het ontwerpen van interventies om mensen te helpen zichzelf te helpen, kan het helpen om emoties te gebruiken om de menselijke cognitieve beperkingen te omzeilen. Wees er daarbij van bewust dat gebeurtenissen meerdere emoties kunnen oproepen, die niet altijd bevorderlijk zijn voor mensen hun doelen. Als vijfde, bij het ontwerpen van een beweegloterij of andersoortige leefstijlloterij, is het aanbevolen om extra aandacht te geven aan designkenmerken die 1) de loterijen aantrekkelijk maken voor mensen om aan mee te doen 2) succesvol blijven na meerdere herhalingen en mogelijk oneindige herhaling van de loterijen en 3) die bijdragen aan de implementatie van de loterijen in de praktijk.

Dankwoord

Ongeveer vier jaar geleden onderbrak ik mijn vakantie om te solliciteren als promovendus. Ik vloog terug uit Spanje voor een eerste gesprek in Tilburg. Gewapend met rolkoffer kwam ik dus vertellen dat het wel een goed idee zou zijn als ik de vacature zou gaan invullen. Tot mijn vreugde bleek - na nog een tweede onderbreking van mijn vakantie - dat de sollicitatiecommissie erin getrapt was, en ik aan de slag kon. Het is haast onvoorstelbaar dat het nu klaar is. Na een turbulente en gedenkwaardige periode, sta ik voor de verdediging van mijn proefschrift. Op deze plaats wil iedereen bedanken die heeft bijgedragen aan de totstandkoming hiervan.

Het team; Mattijs, Jolanda, Johan, Henriëtte en Marcel.
Jolanda, bedankt voor je scherpe pen en tijdige, kritische blik. Je brede kijk op ons werk en het project hielpen me bij de inhoud en de structuur van het werk. Je hebt nooit onderschat hoe waardevol het als promovendus is om snel en uitvoerig feedback te krijgen op werk en ideeën.

Johan, bedankt voor je onvoorwaardelijke vertrouwen en de vele lofzangen. Ik heb genoten en geleerd van je redekunsten en je ontspannen kijk op het werkende leven: “Er is een tijd van zaaien en van oogsten”. Dankzij jou raakte ik steeds verder geïnteresseerd in de gezondheidseconomie. Je neemt altijd de tijd om bij te praten en bent erg toegankelijk. Je optimisme en openheid zijn aanstekelijk.

Henriëtte, jou ben ik ook enorm dankbaar voor het onvoorwaardelijke vertrouwen en je niet aflatende enthousiasme. Onze overleggen waren telkens inspirerend, niet het minst door je visie op wetenschap en de maatschappij. Je hebt van mij een beetje een econoom gemaakt. Je stelde me zonder schroom voor aan je indrukwekkende netwerk en gaf me zo de kans om mijn onderzoek te bespreken met grootheden als Max Bazerman en David Laibson. Dankzij jou bezocht ik Harvard en de Senaat, prachtige ervaringen die ik nooit vergeet.

Marcel, je bent geen officieel onderdeel van het promotieteam, maar ik reken je er toch toe. Ik schreef mijn bachelor scriptie bij jou en jaren later was ik weer welkom om eens te komen praten. Je hebt uiteindelijk belangeloos bij ieder hoofdstuk een belangrijke rol gespeeld. Ik vond het mooi dat overleggen soms voor de helft bestonden uit reisverhalen en anekdotes uit ‘het wereldje’. Je hebt het uitzonderlijke vermogen om complexe zaken helder te verwoorden en hier heb ik enorm veel van geleerd. Het is een eer om met zo’n goede psycholoog samen te werken.

Mattijs, ik kom superlatieven te kort om jou te kunnen bedanken. Ik had me geen betere copromotor kunnen wensen. Je gaf me vertrouwen om mijn eigen pad te bewandelen, maar stuurde subtiel, constructief en tijdig bij. Ik kan je altijd bereiken en ik heb nooit getwijfeld of ik je wel ergens mee zou mogen storen; je hielp áltijd en altijd goed. Dankzij jou heb ik nooit het vertrouwen in dit proefschrift verloren. Onze wekelijkse overleggen gaven me oplossingen en energie. Promoveren is een ontwikkeling tot zelfstandig onderzoeker en ik heb dat mogen

leren van een van de besten. Bedankt voor je meelevendheid, creativiteit, humor, lef en het bieden van fantastische kansen voor mijn toekomst.

Dit proefschrift was er ook niet geweest zonder High Five Health Promotion. Vanaf ons eerste contact was er direct veel interesse in wetenschappelijk onderzoek, wat heeft geleid tot een bijzonder succesvolle en leuke samenwerking. Stef, Paul, Katja, Barry, Irma, Jeannette, Alexander, Merel, Marijn, Stefan, Ruby, Leon, Bas, Bjorn, Ginger en Eric, enorm bedankt voor jullie inzet. Jullie hebben je sportlocaties en sporters vol overgave en zonder enige garantie op succes meer dan een jaar lang beschikbaar gesteld voor dit project, wat uniek en lovenswaardig is.

Ik wil ook de Nederlandse Loterij en betrokken notarissen bedanken voor het verzorgen van de loterijtrekkingen en hulp bij juridische zaken. Geert-Jan, bedankt voor het meedenken en de prettige samenwerking. Ik hoop je nog eens in te halen bij de CPC-run in Den Haag.

Ik wil mijn collega's van Tranzo bedanken voor alle inhoudelijke hulp, maar ook de geslaagde teamuitjes en natuurlijk de heidagen. Met name Hilde, David, Andrea en Arthur; bedankt hiervoor. Ook de collega's van het RIVM, het centrum Voeding Preventie en Zorg, ben ik dankbaar voor de fijne werkomgeving. Ik heb het geluk gehad dat ik de afgelopen jaren onderdeel was van de leukste afdeling van het RIVM: Kwaliteit van Zorg en Gezondheidseconomie. Ondanks de vele commentaren op mijn Brintagebruik zijn jullie heel prettige mensen. Bedankt voor de goede sfeer.

Maarten, Ardine en Claudia, jullie ben ik heel dankbaar voor het meedenken in het vroege ontwerpstadium van de beweegloterij. Dit heeft substantieel bijgedragen aan ons onderzoek. Maarten, heel veel dank ook voor je hulp bij de analyses en natuurlijk de geslaagde cursussen. We zien elkaar in de gym.

Jan Brouwer de Koning, Jaap van Deurzen en Matthijs van den Berg, veel dank voor jullie hulp in het contact met de pers en de succesvolle mediaoptredens. Dames van het secretariaat, vooral Margot, Marianne, Marjan en Truus, bedankt voor jullie ondersteuning bij het project en de bijhorende reizen en bijeenkomsten.

Jongelui van het RIVM, allemaal bedankt voor de gezelligheid tijdens en na het werk. Dat promoveren is soms even zweten (© Johan Polder, 2014) en dan is het fijn als er mensen zijn die je cynisme willen aanhoren (althans doen alsof). Ik heb genoten van de borrels, diners, feestjes, loopwedstrijden, bioscoop- en pretparkbezoeken. Niet uitsluitend, maar in het bijzonder Sander, Yvette, Manon, Francy, Koen, Annerieke, Marjolein en Roy bedankt voor de gezelligheid de afgelopen vier jaar.

Roy en Marjolein, ik ben heel blij met jullie als mijn paranimfen. Marjolein, we hebben een innovatieproject van de grond gekregen in een organisatie met zó'n stoffig imago dat Susan er iedere dag oorverdovend van moet niezen. Onze samenwerking was een heel leuk en gedenkwaardig zijpad van mijn promotie. Bedankt voor je wijze raad, gezelligheid, energie en je inzet in aanloop naar 14 december.
Roy, nooit gedacht dat ik het zo goed zou kunnen vinden met iemand uit Limburg. Vanaf het begin van onze bescheiden carrières hadden we veel gemeen. Het is heel fijn om met iemand te kunnen reflecteren op het leven als promovendus. Je hebt eindeloos veel interesses en hobby's waarmee je me hebt aangestoken. Het is best handig dat iemand alle films en series al gezien heeft, zodat je weet welke je kunt overslaan en welke je moét zien. Bedankt voor je humor, vriendschap en inzet in aanloop naar 14 december.

Op deze plaats wil ik ook de buffels bedanken voor het veroorzaken van veel van de momenten waarop ik niet aan mijn proefschrift dacht. Sander, Lars, Rik, Luuk, Pim, Mark, Tom, Ruben, Max, Maarten, Aart en Lucas, bedankt voor de vele gezellige avonden, zaalvoetbalwedstrijden, stadionbezoeken, bowling- lasergame- en kartevents en natuurlijk de prachtige vakanties naar Namen, Lille en Belgrado. Jullie vriendschap wordt erg gewaardeerd. Zonder jullie had ik de afgelopen vier jaar een stuk minder gelachen dan ik nu heb gedaan.

John, Petra en Ronald, jullie hebben Linda en mij altijd ontzettend uit de brand geholpen, waarvoor eindeloos veel dank. Het is bewonderenswaardig dat jullie zonder enige schroom tijdens twee hectische periodes meehielpen met verhuizen. Jullie passen met plezier op Bounty en verwelkomen ons jaarlijks voor een heerlijk ontspannen vakantie in jullie vakantiehuis. Bedankt dat jullie altijd voor ons klaarstaan en we altijd op jullie kunnen rekenen.

Mijn familie, Marcel, Madeleine, Igor, Ditte, Bas, Ben, Maja, Amélie, Annelou, Janny, Heleen, Eelco, Jeroen, Freija en Stefan, je kunt wel zeggen de we de afgelopen jaren in de hoek zaten waar de klappen vielen. We hebben veel moeten incasseren in korte tijd en ik wil jullie bedanken voor de steun die we daarbij aan elkaar gehad hebben. We bevinden ons nu gelukkig in rustiger vaarwater en ik hoop dat we daar even blijven varen. Uiteraard ook bedankt jullie interesse in mijn werk en de gezellige familieweekends, wandelingen, diners en lunches. Dr. Mous en collega's van de afdelingen hematologie en radiologie van het UMCU wil ik op deze plaats ook oneindig veel bedanken voor hun geneeskunde en uitstekende zorg voor mijn moeder.

Bounty bedankt voor de goede sfeer bij het thuiswerken en de inspirerende lunchwandelingen. Je bijt je altijd volledig vast in de materie. Dankzij jou leverde ik vaak gelikte stukken af. Linda, het leven met een stoffige wetenschapper is niet altijd eenvoudig. Toch heb je geweldig met me meegeleefd en heb je me gelukkig ook helpen relativeren. Je onderbrak mijn hardop denken regelmatig met de liefdevolle boodschap dat het nu gewoon even weekend is. Ik heb de afgelopen jaren onvoorstelbaar veel gehad aan je steun, gezelligheid, liefde en vooral je humor. Ik ben heel blij dat we gaan trouwen.

List of publications

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Schipper, M., Zeelenberg, M., Polder, J. J., & Prast, H. M. (2016). Design and protocol of the weight loss lottery- a cluster randomized trial. *Contemporary Clinical Trials, 49*, 109-115.

Van der Swaluw, K. (2017). The Life Saving Science of Avoiding Temptation. *Harvard Kennedy School Review.* http://ksr.hkspublications.org/2017/10/25/the-life-saving-science-of-avoiding-temptation/.

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Schipper, M., Zeelenberg, M., Berkhout, S., Polder, J.J., & Prast, H. M. (2018). Commitment lotteries promote physical activity among overweight adults—a cluster randomized trial. *Annals of Behavioral Medicine, 52*(4), 342-351.

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Schipper, M., Zeelenberg, M., Berkhout, S., Polder, J.J., & Prast, H. M. (2018). Physical activity after commitment lotteries: examining long-term results in a cluster randomized trial. *Journal of Behavioral Medicine 41*(4), 483-493.

Van der Swaluw, K., Lambooij, M. S., Mathijssen, J. J. P., Zeelenberg, M., Polder, J. J., & Prast, H. M. (2018). Emotional responses to behavioral economic incentives for health behavior change. *Psychology, Health & Medicine 23*(8), 996-1005.

Struijs, J., Hayen, A., & **Van der Swaluw, K.** (2018). When Designing Bundled Payments, Don't Ignore The Lessons Of Behavioral Economics. *Health Affairs blog.* https://www.healthaffairs.org/do/10.1377/hblog20180420.640240/full/.

About the author

Koen van der Swaluw was born on 30 September 1990 in Wageningen (the Netherlands). After completing his Atheneum at RSG Pantarijn in 2009, he moved to Tilburg to study Psychology. He received his Bachelor's degree in Psychology from Marcel Zeelenberg at Tilburg University in 2013. In 2013, Koen moved to Utrecht to start the Master Social Psychology at Utrecht University. He graduated in 2014, advised by Madelijn Strick.

After his studies, Koen van der Swaluw started his PhD project about lifestyle behavior change at Tilburg University (Tranzo Scientific Center for Care and Welfare) and the National Institute for Public Health and the Environment (department of Health Economics) in 2014, advised by Johan Polder, Henriette Prast, Mattijs Lambooij and Jolanda Mathijssen. After completing his PhD project in 2018, Koen remains working on the understanding and facilitation of behavior change in the fields of health, sustainability and safety.