



HAL
open science

Is There a Gender Gap in the Birthday-Number Effect? The Case of Lotto Players and the Role of Sequential Choice

Catherine D Hondt, Patrick Roger, Arvid O I Hoffmann, Daria Plotkina

► **To cite this version:**

Catherine D Hondt, Patrick Roger, Arvid O I Hoffmann, Daria Plotkina. Is There a Gender Gap in the Birthday-Number Effect? The Case of Lotto Players and the Role of Sequential Choice. 2023. hal-04451848

HAL Id: hal-04451848

<https://hal.science/hal-04451848>

Submitted on 12 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Is There a Gender Gap in the Birthday-Number Effect?

The Case of Lotto Players and the Role of Sequential Choice

Catherine D'Hondt,^{†1} Patrick Roger,* Arvid O. I. Hoffmann,[§] Daria Plotkina[‡]

Abstract

The literature on lottery gambling shows that players do not select numbers randomly, a phenomenon which is called *conscious selection*. Mainly, players prefer “small” numbers (less than thirty), either because of the existence of small lucky numbers or because they are victims of the so-called *birthday-number effect*. Because lotto games are parimutuel, such preferences result in poor ticket choices in terms of achieving below average returns. Using data from Belgium, where approximately 10% of the population plays lotto games every week, this paper extends prior literature by documenting the existence of a gender gap in the birthday-number effect, with women displaying a stronger birthday-number effect than men, as well as the non-persistence of the birthday-number effect (and consecutively of the gender gap) when participants are asked to fill in a second lotto ticket immediately after their first one. The disappearance of the birthday-number effect in sequential choices appears to be driven by response speed, with participants being twice as fast to fill in the second ticket compared to the first one. Moreover, we find that participants who bet on their birthday numbers take significantly more time to complete their ticket. Contrary to prior research, we find that the strength of the birthday-number effect is positively related to deliberative number choices, not intuitive and automatic number choices. Our results are robust to controlling for potential confounding effects including those related to participants' age, education, self-esteem, and superstitious beliefs.

Keywords: Birthday-number effect; gender gap; conscious selection; sequential choices; lotteries.

Declaration of interests: The authors have no relevant financial or non-financial interests to disclose. No funding was received for conducting this study.

Ethics approval and consent: All procedures performed in studies involving human participants were in accordance with the ethical standards of the research institutions and did not require a separate ethics committee approval. This study was conducted from publicly available data and survey data with informed consent.

Data: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authorship contribution statement:

Catherine D'Hondt: Conceptualization, Methodology, Formal analysis, Writing – original draft.

Patrick Roger: Conceptualization, Methodology, Formal analysis, Writing – original draft.

Arvid Hoffmann: Conceptualization, Methodology, Writing – review and editing.

Daria Plotkina: Conceptualization, Methodology, Writing – review and editing.

^{†1}Louvain Finance (LIDAM), Louvain School of Management, UCLouvain – Mailing address: Chaussée de Binche 151, 7000 Mons, BELGIUM – Tel: +32 (0)65323339 – Email: catherine.dhondt@uclouvain.be ; *Corresponding Author*

*LaRJE Research Center, IAE NC, University of New Caledonia _ Mailing address: IAE NC, Nouville, 61 Cedex FRANCE – Tel: +687 290364 – Email: patrick.roger@unc.nc .

[§]University of Adelaide Business School – Mailing address: 10, Pulteney Street, Adelaide, South Australia, SA5005, Australia – Tel: +61 883134223 – Email: arvid.hoffmann@adelaide.edu.au.

[‡]EM Strasbourg Business School – Mailing address: 61 avenue de la Forêt Noire, 67085 Strasbourg Cedex FRANCE – Tel: +33 (0)368858252 – Email: daria.plotkina@em-strasbourg.eu.

Is There a Gender Gap in the Birthday-Number Effect?

The Case of Lotto Players and the Role of Sequential Choice

Abstract

The literature on lottery gambling shows that players do not select numbers randomly, a phenomenon which is called *conscious selection*. Mainly, players prefer “small” numbers (less than thirty), either because of the existence of small lucky numbers or because they are victims of the so-called *birthday-number effect*. Because lotto games are parimutuel, such preferences result in poor ticket choices in terms of achieving below average returns. Using data from Belgium, where approximately 10% of the population plays lotto games every week, this paper extends prior literature by documenting the existence of a gender gap in the birthday-number effect, with women displaying a stronger birthday-number effect than men, as well as the non-persistence of the birthday-number effect (and consecutively of the gender gap) when participants are asked to fill in a second lotto ticket immediately after their first one. The disappearance of the birthday-number effect in sequential choices appears to be driven by response speed, with participants being twice as fast to fill in the second ticket compared to the first one. Moreover, we find that participants who bet on their birthday numbers take significantly more time to complete their ticket. Contrary to prior research, we find that the strength of the birthday-number effect is positively related to deliberative number choices, not intuitive and automatic number choices. Our results are robust to controlling for potential confounding effects including those related to participants’ age, education, self-esteem, and superstitious beliefs.

1. Introduction

Lotto games are popular all around the world.² Playing the game is simple; requiring only the selection of a subset of n numbers out of m ($n \ll m$).³ As lotto games are parimutuel, the global amount of money redistributed to winners does not depend on the number of winners but only on the number of tickets sold. Betting on popular numbers (whatever the reason for their popularity) is therefore detrimental, because prizes are shared equally among a larger number of winners, leading to lower individual prizes. Betting on unpopular numbers improves the expected return, though it remains negative due to the takeout rate (close to 50% in most countries).

A large body of research on lottery gambling shows that players do not select numbers randomly, a widespread phenomenon coined as *conscious selection*. Not only numbers perceived as “lucky” are popular and thus overplayed in lotto games. Other potential reasons explaining suboptimal number choice by players include the *birthday-number effect*. In short, this effect involves people betting on numbers representing their own birthday or those of relatives, with the consequence that numbers below 30 are overplayed and therefore less profitable than numbers greater than 31 (Baker and McHale, 2009, 2011; Cook and Clotfelter, 1993; Farrell et al., 2000; Finkelstein, 1995; Roger and Broihanne, 2007; Simon, 1999; Wang et al., 2016; Roger et al., 2023).

Although a lot of research in economics has been devoted to understanding why people bet on unfair lotteries (Rogers, 2008; Stetzka and Winter, 2021), not much has been written to identify the individual drivers of the birthday-number effect (or more generally of conscious selection). Baker and McHale (2011) suggest a hierarchy of number preferences, with “lucky” numbers at the top (e.g., the number 7 in western countries; the number 8 in China), followed

² According to Globe Newswire, the world lottery sales amounted to \$211 billion in 2022 (<https://www.globenewswire.com/news-release/2023/06/26/2694785/0/en/Lottery-Global-Market-Report-2023.html>).

³ For the Belgian lotto considered in this paper, $n = 6$ and $m = 45$.

by “easy-to-remember” numbers, mainly birthdays and anniversaries. Husemann-Kopetzky and Kocher (2016) also identify number preferences among supporters of sports teams with numbers in the club’s name (e.g., the German soccer clubs Schalke04 or BVB09 (Borussia Dortmund)). Polin et al. (2021) link the popularity of numbers in the Israeli lotto game to their location on the form. These authors conclude that people prefer numbers located in the first three rows (out of four) of the form, corresponding to numbers below 27. They also find that conscious selection decreases when sales increase, with popular numbers becoming less popular and unpopular numbers becoming less unpopular. Wang et al. (2016) find that people favor personally and situationally meaningful numbers. Roger and Broihanne (2007) find that the most popular numbers in the French 6/49 lotto (over a 25-year period) are 7, 12, 9, 13, 11, 5 and the least popular are 32, 41, 39, 40, 38, 43. Recently, Roger et al. (2023) find 7, 12, 13, 17, 18, 28 (35, 40, 43, 44, 45) as the most (least) popular numbers in the 6/45 Belgian lotto over a 4-year period from 2017 to 2021.

None of the abovementioned studies addresses the question of the individual drivers of the birthday-number effect in lottery gambling as they were merely based on an analysis of tickets played, without any background information on their players. To the best of our knowledge, the only study briefly raising the question of a potential gender gap in the birthday-number effect is Kitayama and Karasawa (1997). Focusing on self-esteem⁴ in Japan, these authors examine the birthday-number effect as a complement to the so-called *name-letter effect*, which reflects people preferring letters of their own name (Nuttin, 1985). For the birthday-number effect, Kitayama and Karasawa (1997) asked students how much they like any given number using a six-point Likert scale, showing that the average score is significantly higher for birthday compared to non-birthday numbers, and that women seem more prone to display this preference. Given the purpose of their research, Kitayama and Karasawa (1997) did not rely on

⁴ Kitayama and Karasawa (1997) refer to implicit self-esteem, defined as the introspectively unidentified effect of the self-attitude on evaluation of self-associated and self-dissociated objects.

a lottery gambling setting and therefore provide no insights into the birthday-number effect in the context of lotto number choices. Moreover, these authors only report univariate results consistent with a gender gap, meaning that they cannot rule out potential confounding effects due to other individual characteristics. Finally, Kitayama and Karasawa (1997) did not examine persistence of a potential gender gap in the birthday-number effect across sequential choices.

This paper is an attempt to fill the gap in our understanding left by Kitayama and Karasawa (1997) by investigating the existence of a gender gap in the birthday-number effect among lotto players. For that purpose, we developed and administered a dedicated survey among 506 Belgian individuals. We asked participants to fill in two lotto tickets in a row, allowing us to both measure the birthday-number effect and check its persistence. Later in the questionnaire, we asked their degree of agreement (on a seven-point Likert scale) with the statement “When I play lotto, I select numbers linked to my birthday.” We therefore got a direct measure of the birthday-number effect through participants’ lotto ticket choice (revealed preference) and an indirect one with participants’ score on the Likert scale (stated preference).

Using the aforementioned two sets of information enables us to contribute to the literature in several ways. First, we document the existence of a birthday-number effect in the number choices of lotto players, beyond the observation of conscious selection and preference for small numbers previously reported in the literature. Second, we confirm that the birthday-number effect is stronger among women, even after controlling for usual socio-demographic factors such as age and education, and psychological traits like self-esteem and superstitious beliefs. Third, we provide evidence showing that the birthday-number effect is not persistent when people are asked to fill in two tickets in a row. Indeed, the numbers selected on the second lotto ticket demonstrate the disappearance of the birthday-number effect (and the associated gender gap) in sequential choices, though conscious selection is still at work (because participants continue to prefer small numbers, even when filling in their second lotto ticket).

Of particular interest, people in our sample selected numbers “more randomly” on the second ticket, mainly because they selected numbers more quickly, without thinking long before notching down numbers. This significantly faster decision process of number selection for the second ticket suggests that people were selecting numbers in a more automatic or intuitive manner, leading to choices closer to a random draw. By contrast, participants who bet on their birthday numbers took significantly more time to complete their ticket, indicating more deliberative number choices. Contrary to Koole et al. (2001) who focus on self-esteem⁵ and report that the positive bias for birthday numbers is inhibited when participants are induced to respond in a deliberative manner,⁶ we provide evidence supporting that the birthday-number effect is larger when players take more time to select numbers, suggesting that playing birthday numbers is the result of a deliberative choice, not of an intuitive and automatic choice. Furthermore, our result showing that the birthday-number effect vanishes on the second lotto tickets extends those of Wang et al. (2016) who find that conscious selection (non-random choice) weakens when sales increase, that is, when people increase the amount they bet and buy multiple tickets.

The remainder of this paper is organized as follows. Section 2 presents our testable hypotheses. Section 3 describes our survey data and provides descriptive statistics about the sample. In Section 4, we discuss the representativeness of our survey data and provide univariate results. Multivariate results are reported in Section 5. Section 6 concludes the paper.

2. Hypotheses

In gambling, conscious selection essentially means excessive preference for small numbers (less or equal to 30). As mentioned earlier, several reasons have been put forward in the

⁵ Like Kitayama and Karasawa (1997), Koole et al. (2001) use overevaluation of name letters and birthday numbers as measures of implicit self-esteem. However, these authors mainly examine the impact of contextual factors on self-esteem, and not its individual drivers.

⁶ In their Study 2, Koole et al. (2001) split their sample in such a way that half participants were specifically asked to think to their letter preferences before providing preference grades to alphabet letters.

literature to explain this widespread tendency. The most frequently cited are the existence of small lucky and/or hot numbers (Rogers, 1998), the availability of “easy-to-remember” numbers like birthdays and anniversaries (Wang et al., 2016), or the tendency to select numbers located at the top of the lotto grid (Polin et al., 2021). Since this paper focuses on the birthday-number effect, we formulate our first hypothesis as follows:

H1: *Participants tend to bet on their birthday numbers (or those of relatives) when selecting numbers on a lotto ticket.*

Building on Kitayama and Karasawa (1997), who report that women are more prone to display a positive bias for their birthday numbers than men, we expect to find a gender gap in the birthday-number effect. Such an expectation is also consistent with the results of Auria (1993) showing that women are better than men in remembering dates, especially when dates are self-related. Hence, we posit our second hypothesis as follows:

H2: *The birthday-number effect in the selection of lotto numbers is stronger among women than men.*

The birthday-number effect (alongside with the name-letter effect) has been documented as a valid measure of self-esteem (Kitayama and Karasawa, 1997; Koole et al., 2001). Therefore, we conjecture self-esteem could be a driver of the heterogeneity of the birthday-number effect among lotto players, leading to the following testable hypothesis:

H3: *Participants with higher levels of self-esteem are more likely to be subject to the birthday-number effect than their counterparts.*

Another psychological trait that might play a role in the birthday-number effect, and possibly in the existence of a gender gap, is the level of superstitious beliefs that is recognized as a major individual determinant of gambling activities. According to Rogers (1998), superstitious beliefs reflect people’s failure to appreciate the lack of a cause-and-effect relationship between their gambling behavior and subsequent events in the world around them.

More specifically, past research shows that the level of superstitious beliefs is an important factor influencing lottery players' selection of specific numbers (Chou et al., 2009; He et al., 2020; Hirshleifer et al., 2018; Roger et al., 2023). Regarding the birthday-number effect, we posit that lotto players with higher superstitious beliefs are more likely to consider their birthday numbers (or those of their relatives) as lucky numbers, compared to their counterparts. As a result, this tendency should lead them to display a stronger birthday-number effect than lotto players with lower superstitious beliefs. Our next hypothesis is therefore stated as follows:

H4: *Participants with higher levels of superstitious beliefs are more likely to be subject to the birthday-number effect than their less superstitious counterparts.*

Finally, building on Koole et al. (2001), we conjecture that selecting numbers for a single lotto ticket or selecting numbers for several tickets in a row could lead to different intensities in the birthday-number effect. Koole et al. (2001) show that asking participants to think about why they prefer some numbers leads to the disappearance of the preference for birthday dates. In our setting where participants fill in two tickets in a row, we wonder whether a mitigation of the birthday-number effect is likely to appear on the second ticket. There are at least two reasons for raising that question. On the one hand, people might start thinking when filling in the second ticket after having automatically selected (easy-to-remember or lucky or birthday) numbers on the first ticket. On the other hand, they might change most of the selected numbers on the second ticket to opt for a more random-looking sequence in an automatic manner. This latter possibility would be consistent with the findings of Wang et al. (2016) showing that popular (unpopular) numbers become less popular (unpopular) when people play more tickets. Considering these two possible phenomena, we formulate our last hypothesis as follows:

H5: *The strength of the birthday-number effect decreases when participants fill in the second lotto ticket, compared to the first ticket.*

3. Data

3.1 Data collection

In order to investigate the birthday-number effect and test our hypotheses, we recruited 506 Belgian participants for an online survey. Participants were part of a panel of individuals managed by the data provider Qualtrics, using quotas to reflect national statistics regarding socio-demographics (i.e., age, gender, income, and education) as closely as possible. As Belgium has a French-speaking part (Wallonia) and a Dutch-speaking part (Flanders), participants could answer either in French or in Dutch, with the sampling strategy aiming to get an approximate 50-50 split of participants between both parts.

We excluded 8 participants with incomplete or invalid responses. The remaining 498 participants were asked to fill in two 6/45 lotto tickets in a row and respond to a set of socio-demographic and psychographic questions. More precisely, at the beginning of the questionnaire, participants were asked to provide their age, gender, language (Dutch or French), marital status, number of children below 18, education level, region of living, postal code, income, frequency of lotto playing during the last year and use of the Quick Pick system.⁷ It is important to note that no question about birthday numbers was asked before participants completed the two lotto tickets.

3.2 Sample description

The final sample of $N=498$ participants was almost equally distributed between men and women ($n_{\text{men}}=260$ and $n_{\text{women}}=238$) and participants from Wallonia and Flanders ($n_{\text{Wallonia}}=286$ and $n_{\text{Flanders}}=212$). The average age of the participants was 45.13 years ($SD=15.8$), and the majority has completed secondary school (54.22%), while the rest held university degrees (25.5% had a bachelor's degree and 19.3% had a master's degree). Within the sample, 36.95% were married, 24.3% were single, and 11.24% were divorced. On average, the participants had

⁷ This tool allows an automatic and random selection of lotto numbers.

1.52 underage children, 36.5% lived in a rural, 20% in a suburban, and 43.5% in an urban area. Over the last year, 20% of participants played lotto at least once a week, 15% several times a week, while 28% played from multiple times to once a month, and 37% played rarely to never. In our empirical work (see Sections 4 and 5), we considered the fact that regular players and occasional players or non-players might behave differently.

3.3 Variables and measurement scales

We measured the propensity to play birthday numbers in two ways. On the one hand, to get an indirect measure of the birthday-number effect (stated preference), we asked participants to assess on a seven-point Likert scale their agreement (1 = fully disagree, 7 = fully agree) with a statement saying that they would select their own birthday numbers when playing lotto. We call this variable *Birthday_Numbers*. In addition, we asked a similar question addressing playing the birthday numbers of other people (e.g., relatives, friends, children), which provides us with a variable named *Birthday_Others*. We also asked similar questions delivering stated preference about playing numbers related to one's age and playing lucky numbers.

On the other hand, to get a direct measure of the birthday-number effect (revealed preference), we extracted the actual propensity of participants to play their birthday numbers using the six numbers selected on their lotto tickets. It is worth recalling that participants had to select these 6 numbers from 1 to 45 on two consecutive lotto tickets whose template corresponds to the one of the national Belgian 6/45 lotto. Following a two-step approach for convenience, we first defined two dummy variables as follows: the variable *Birthday_Day* is set to 1 for participants who selected their birthday day, while the variable *Birthday_Month* is set to 1 for participants who selected their birthday month. Next, we defined another dummy variable, called *Play_Date*, which is set as the maximum value of *Birthday_Day* and *Birthday_Month*. Put differently, *Play_Date* is equal to 1 when participants selected at least either their birthday day or their birthday month on the lotto ticket. Note that all these direct

measures of the birthday-number effect are determined separately for each of the two lotto tickets filled in by participants.

For the psychographic variables, we used established multiple-item scales, all measured on seven-point Likert scales for consistency. To evaluate participants' superstitious beliefs in terms of a personality trait (see H4), we used four items from Carlson et al. (2009), while their self-esteem (see H3) was measured with ten items of Rosenberg's (1965) Self-Esteem Scale. We validated the superstition scale reliability with Cronbach's alpha value being higher than 0.79 and that of self-esteem being higher than 0.88 (Netemeyer et al., 2003).

3.4 Common method variance

Considering we used a single online survey to collect all of our data, there is the potential risk of common method variance (CMV) bias influencing our results. We tested for this bias following previously established statistical procedures (Craighead et al., 2011). First, a Harman's single-factor test with an exploratory factor analysis confirmed that the different measured variables did not all load onto a single factor (Podsakoff et al., 2003). Next, a marker variable test confirmed that a theoretically unrelated question on participants' attitude towards the color blue (Simmering et al., 2015) did not correlate with the key psychographic and socio-demographic variables. All these tests show that CMV bias is not a concern for our results.

4. Survey data representativeness and univariate results

To provide model-free evidence before showing the results of a regression analysis, we will show in Section 4.1 that i) our survey data are consistent with field data in terms of conscious selection, and ii) conscious selection manifests itself by a disproportionate choice of numbers under 30, meaning that the non-random choice of numbers might be linked to the birthday-number effect. In Section 4.2, we will show that individuals in our sample iii) declared a tendency to select birthday numbers, and iv) selected their birthday numbers significantly more than what is

expected under a random choice assumption. In Section 4.3, we will also provide evidence showing that women are, on average, more prone to display the birthday-number effect than men.

4.1 Representativeness

Figure I-A shows two different curves; a dashed curve based on Belgian lotto field data that is publicly available⁸ and a solid curve based on our survey data. The dashed curve gives the popularity index of single numbers (properly standardized to be comparable with raw frequencies measured in the survey), calculated from Belgian lotto data for the period from May 26, 2018 to December 31, 2022, corresponding to 481 draws. The starting date is selected for the sake of homogeneity. It corresponds to the first draw with 9 ranks of gain and a ticket price at 1.25€. Before that date, there were only 8 winning ranks, and the ticket price was 1€. The popularity index is calculated following Roger et al. (2023), and then standardized in such a way that the vector of popularities has the same expectation and variance as the vector of raw frequencies, for the sake of comparability. In short, the popularity index of a number k is defined as the average proportion of winners over ranks 1 and 3 when the number k under consideration is on the winning ticket (details are in Appendix A.4 of Roger et al. (2023)). The solid curve is the frequency of number choices over the 498 first lotto tickets selected by participants in our survey sample. The two horizontal lines define the 99% confidence interval of frequencies under a random choice assumption. Under random choice, the frequency of a given number follows a binomial distribution with parameters $N = 498$ and $p = 6/45$ that can be approximated by a Gaussian distribution with an expectation equal to Np and a variance equal to $Np(1-p)$. It is therefore easy to calculate the confidence interval outside of which numbers could be said (un)popular at the given significance level (99% on Figure 1-A).

Both curves in Figure I-A are in line with prior papers on the topic (Farell et al., 2000; Polin et al., 2021; Roger and Broihanne, 2007; Roger et al., 2023; Wang et al., 2016). That is,

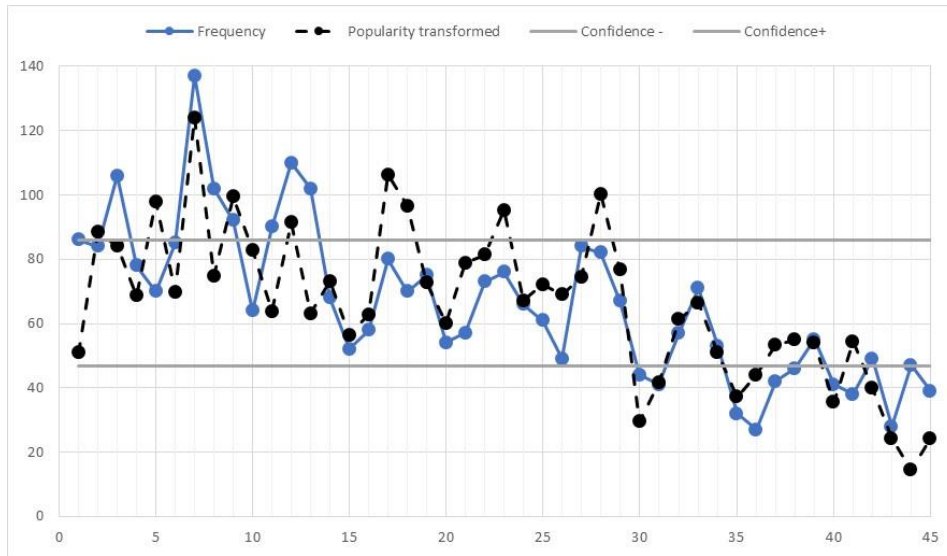
⁸ Time series of Belgian lotto results can be freely downloaded at <https://www.loterie-nationale.be/nos-jeux/lotto/resultats-tirage/statistiques>.

there is a strong preference of lotto players for small numbers compared to a dislike for numbers greater than 30. In fact, numbers that might correspond to a month and/or a day in a birthdate are preferred to numbers in the last third of the range (since the Belgian lotto is a 6/45 game with one bonus number). The similarity of the two curves indicates that the survey data used in this study are relevant to analyzing conscious selection and the birthday-number effect since the number choices by our survey participants replicate the stylized facts observed in past research using publicly available field data from the Belgian lotto.

Moreover, the shape of both curves seems relatively stable over time because Figure 4 in Roger et al. (2023) as well as Figure 3 in Polin et al (2021) are obviously similar to Figure I-A with different survey data and different time periods for field data. To provide further evidence, Figure I-B shows the cumulative distribution function (CDF) of number choices in our sample. The observed CDF is always above the straight line corresponding to the theoretical CDF of a random choice (in case of a random choice, we should get the same proportion ($1/45$) for each of the 45 numbers). The first-order stochastic dominance of the theoretical CDF over the observed CDF confirms the tendency of participants to choose small numbers. Finally, conscious selection, and possibly the birthday-number effect, appears clearly in Figure I-C that represents the difference between the observed CDF and the theoretical CDF under random choice. The curve decreases sharply over the last third part, that is, for numbers higher than 30, showing again the lower popularity of large numbers, compared to numbers below 30.

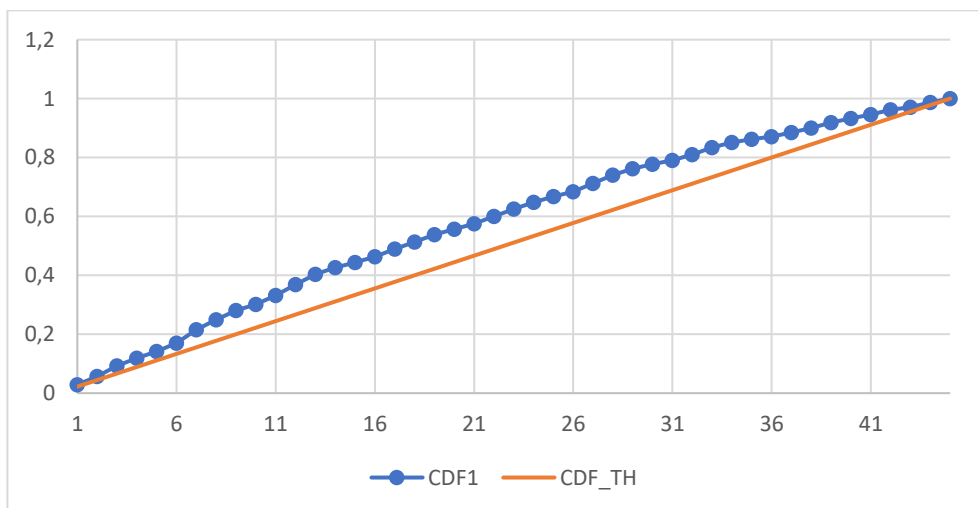
Figure I: Aggregate choices ($m \times n = 498 \times 6$ numbers for each draw)

Figure I-A: Frequencies in survey data and standardized popularity from field data



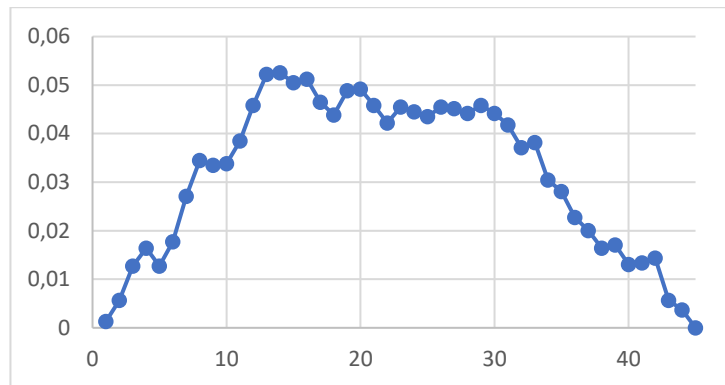
Note: The horizontal axis reports the 45 numbers. The dashed curve gives the standardized popularity index of numbers, calculated from publicly available Belgian lotto data over the period 2018-26-05 to 2022-31-12 (corresponding to 481 draws). The popularity index is calculated following Appendix A.4 of Roger et al. (2023), and then standardized (to deliver a sequence of numbers with the same mean and standard deviation as the sequence of frequencies measured in the survey). The solid curve represents the frequency of number choices over the 498 first lotto ticket choices in our sample.

Figure I-B: Cumulated frequencies of number choices in survey data (in %)



Note: The horizontal axis reports the 45 numbers, and the vertical axis gives the cumulative distribution function of numbers selected in the 498 first lotto tickets in our sample (dotted line). For the sake of comparison, the straight line represents the cumulative distribution function under random choice.

Figure I-C: Difference between observed and theoretical cumulated frequency (in %) in survey data



Note: The horizontal axis reports the 45 numbers, and the vertical axis refers to the difference between the cumulative distribution function of numbers selected in the 498 first lotto tickets in our sample and the corresponding theoretical cumulative distribution function (i.e., the cumulative distribution function under random choice).

Statistical confirmation for the graphical patterns observed in Figure I may be obtained through different univariate tests. First, the sample average of numbers selected on the first lotto tickets (variable *Mean_Draw*) is 19.494, which is significantly below 23, which is the expected value under random choice in a 6/45 game ($t(497) = -15.33, p < .0001$). We also performed a more direct test of the preference for numbers below 30 by defining a variable *Less30* which counts the selected numbers lower or equal to 30 on each lotto ticket (participants selected 6 different numbers on each ticket). Under random choice, the expected value for this variable is equal to 4 in a 6/45 game. Our sample average for *Less30* on the first lotto tickets is 4.663, significantly above the expected value for this variable under random choice ($t(497) = 15.09, p < .0001$).

All in all, our survey data seems representative in terms of depicting the stylized facts observed in field data, which include conscious selection, and suggests that the birthday-number effect might be at work when participants select lotto numbers. However, overplaying numbers below 30 does not always mean playing birthday numbers. The next subsection provides further univariate results testing whether the birthday-number effect is indeed at play.

4.2 Self-declared and direct measures of the birthday-number effect

As explained in Section 3.3, we measured the propensity to play birthday numbers in two ways. Using the indirect measure of the birthday-number effect (stated preference), *Birthday_Numbers*, the average score of participants is significantly higher than 4 at 4.835 ($t(497)=9.42, p<.0001$).⁹ Similarly, the average value of the *Birthday_Others* variable (stated preference for playing birthday numbers of relatives) significantly exceeds 4 with a value of 4.775 ($t(497)=8.50, p<.0001$). The very high correlation (0.802) between *Birthday_Numbers* and *Birthday_Others* is noteworthy. That is, participants who declared a propensity to bet on numbers representing their own birthday also reported a tendency to bet on numbers representing the birthday of other people. Nevertheless, assessing the propensity to bet on birthday numbers using a comparison to the midpoint of the scale might be misleading if participants had a general tendency to use the top grades of the scale, regardless of the question under consideration. However, this is not the case since regarding the questions about playing age numbers and lucky numbers, the degrees of agreement are significantly below 4 at 3.659 for age numbers ($t(497)=-3.73, p=0.0002$) and 3.602 for lucky numbers ($t(497)=-4.43, p<.0001$).

Next, using the two binary variables *Birthday_Day* and *Birthday_Month* as direct measures of the propensity of participants to play their birthday numbers on the first lotto ticket, we can compare their actual values to the expected values under random choice (i.e., if there is no conscious selection). Under random choice, the probability for participants to select their birthdate day (month) on a ticket is $q = 1 - \frac{\binom{44}{6}}{\binom{45}{6}} = 0.1333 = 13.33\%$. For comparison, the sample average of *Birthday_Day* is 59.44% ($t(497)=20.93, p<.0001$) and the sample average of *Birthday_Month* is 39.76% ($t(497)=12.04, p<.0001$), both significantly higher than q .

⁹ We refer to 4 as the midpoint of the seven-point Likert scale.

Hence, the preference for birthday numbers is unambiguous when participants fill in their first lotto ticket. We thus find evidence supporting H1.

Building on Koole et al. (2001), who find that the positive bias for birthday numbers is inhibited when participants are induced to respond in a deliberative manner, we also measured the time (in seconds) used by participants to fill in their first lotto ticket (variable *Lotto_Duration*). To assess whether playing birthday numbers follows a more *reflective* or a more *intuitive* (fast and automatic) mental path, we used the dummy variable *Play_Date* for the first lotto tickets (set to 1 when participants selected at least their birthdate day or month) to split the sample and check for a significant difference in *Lotto_Duration* between the two groups. Participants who bet on their birthday numbers (day and/or month) took on average 34.48 seconds to fill in their first ticket, while those who did not took only 26.79 seconds. This difference, which is meaningful (about 8 seconds) and highly significant ($t(496)=-2.67$, $p=0.0078$), suggests that the birthday-number effect relies on a more reflective mental path.

4.3 Gender and the birthday-number effect

Table 1 provides information about the distribution of birthdate days in the sample, with a distinction between men and women. For the sake of brevity, we sorted the 31 possible days into three ranges (the last one including 11 days, from 21 to 31). Table 1 also reports the proportion of men and women who played their birthdate day on their first lotto ticket in each of these three ranges. Altogether, 63.45% of women selected their birthdate day against 55.77% of men. The difference between these proportions is significant at the 10% level ($t(496)=-1.74$, $p=0.0817$). Considering the three ranges of days, discrepancies are noticeable. Only the first range (i.e., 10 first days) displays a significant difference ($t(173)=-2.29$, $p=0.0235$) between men and women, with men playing their birthday numbers less often than women. In the last range (i.e., 11 last days), the proportion of women playing their birthday numbers is also higher,

but the difference with the proportion of men is not statistically significant at a conventional level. By contrast, in the medium range of days, the proportion of women playing their birthday numbers is somewhat lower than the proportion of men, although the difference is not significant.

Table 1
Distribution of birthdate days and playing behavior among men and women

Range of birthdate days	Number of Men	Number of Women	Number of Men playing their birthdate day	Number of Women playing their birthdate day	Men playing their birthdate day (%)	Women playing their birthdate day (%)	p-value
1-10	91	84	43	54	47.25%	64.29%	0.0235
11-20	92	69	55	39	59.78%	56.52%	0.6801
21-31	77	85	47	58	61.04%	68.24%	0.3412
Total/Average	260	238	145	151	55.77%	63.45%	0.0817

To further examine heterogeneity in the strength of the birthday-number effect among men and women, we tested the existence of a gender gap using the following variables: *Mean_Draw*, *Less30*, *Birthday_Numbers*, *Birthday_Day*, and *Birthday_Month*. Among them, the first two variables measure the tendency to select small numbers, while the three others specifically capture the birthday-number effect. Table 2 reports the results of univariate two sample t-tests comparing the means of the aforementioned variables across genders. Significant differences between men and women are noticeable for all the variables, except for *Birthday_Month*, where we found the expected sign, but the difference was not statistically significant at conventional levels. Unsurprisingly, *Mean_Draw* and *Less30* exhibit consistent results. On average, women selected significantly smaller numbers on their first lotto ticket compared to men, with the average of the selected numbers being 19.92 for men and 19.03 for women ($t(496)=1.95, p=0.0517$). On average, women selected more numbers under 30 on their first lotto ticket than men with an average score of 4.58 for men and 4.76 for women ($t(496)=-$

2.05, $p=0.0412$). Next, women also agreed to a greater extent with the self-assessment question about one's tendency to select birthday numbers when playing lotto with an average score of 4.54 for men and 5.16 for women ($t(495.64)=-3.56, p=0.0004$). As expected, the average value of the dummy variable *Birthday_Day* is also lower for men (0.56), compared to women (0.63), although the degree of significance for this difference is lower ($t(496)=-1.74, p=0.0817$). All in all, the univariate results reported in Table 2 provide support to the existence of a gender gap in the birthday-number effect. However, it is worth mentioning that the results related to *Mean_Draw* are not sufficient to conclude to the existence of a gender gap in the birthday-number effect because women have, on average, their birthday later in the month, as illustrated in Table 1. The gender gap is mainly confirmed by the differences in the other variables, as we will show in the next section.

Table 2
Univariate analysis of gender gap in the birthday-number effect
(first lotto tickets)

Variable	Men	Women	t-value	p-value
<i>N</i>	260	238		
<i>Mean_Draw</i>	19.92	19.03	1.95	0.0517*
<i>Less30</i>	4.58	4.76	-2.05	0.0412**
<i>Birthday_Numbers</i>	4.54	5.16	-3.56	0.0004***
<i>Birthday_Others</i>	4.41	5.18	-4.31	<.0001***
<i>Birthday_Day</i>	0.56	0.63	-1.74	0.0817*
<i>Birthday_Month</i>	0.38	0.42	-0.80	0.4238

Note: This table reports cross-sectional means and proportions among men and women as well as the results of statistical tests comparing these means and proportions. *Mean_Draw* is the average of the numbers selected by participants on their first lotto ticket. *Less30* counts participants' selected numbers lower or equal to 30 on their first lotto ticket. *Birthday_Numbers* refers to the level selected on a seven-point Likert scale assessing participants' self-declared tendency to select their own birthdate numbers when playing lotto. *Birthday_Others* refers to the level selected on a seven-point Likert scale assessing participants' self-declared tendency to select birthday numbers of other people when playing lotto. *Birthday_Day* and *Birthday_Month* are dummy variables set to 1 when participants selected either their birthdate day or month on their first lotto ticket. *N* gives the number of participants. Significance levels: *** $p < .01$, ** $p < .05$, * $p < .10$.

Table 2 does not distinguish regular players (313 participants) who play at least once a month (up to several times per week) from occasional (less than once a month) or non-players (185 participants). Therefore, one could ask whether these results, supposed to reveal differences in male and female players' behavior, only reveal intentions of occasional players (80 participants) or non-players (105 participants). We checked and it turns out that it is not the case for the most important variables *Birthday_Numbers* and *Birthday_Others*. Table A2 in the appendix reports the separate results for regular players and occasional or non-players. Nevertheless, Table A2 shows that among regular players there are no significant differences between men and women concerning the first two variables *Mean_Draw* and *Less30*. Table A2 also reveals that almost no difference appears between female regular players and female occasional or non-players. On the contrary, men who are occasional or non-players behave differently from male regular players. The average value of *Birthday_Numbers* differs substantially between regular players (4.79) and occasional or non-players (3.89) for men only ($t = 3.23, p < 0.01$). For women, the same difference is not significant ($t = 0.63, p = 0.5307$). Moreover, the subsample of male occasional players does not exhibit a preference for birthday numbers in the sense that the average of this variable on this subsample of 73 men is not significantly different from the middle of the scale, namely 4 ($t = -0.42, p = 0.6763$).

5. Multivariate results and persistence of the birthday-number effect in sequential choices

Section 4 provided evidence pointing to the existence of a gender gap in the birthday-number effect. However, we need to control for potential confounding factors before we can validate with certainty H2 about this gender gap. First, being single or in a couple and having children or not are both socio-demographic indicators likely to influence the propensity of people to select birthday numbers of relatives. Second, as posited in H4, having more pronounced superstitious beliefs is likely to make people perceive birthdates as “lucky” numbers. This

individual trait needs to be controlled for since women in our sample perceived themselves as significantly more superstitious than men (average score of 3.78 for women compared to 3.48 for men ($t(496)=-2.32, p=0.0209$)). In the same vein, self-esteem might also play a role in people's tendency to select their own birthday numbers, as stated in H3. In the sample, men rated themselves significantly higher on the self-esteem scale with an average score of 4.96 for men versus 4.79 for women ($t(454.08)=1.72, p=0.0870$). Furthermore, rationality and the ability to understand the functioning of parimutuel games might have prevented participants from playing birthdates because of the loss of expected returns generated by selecting lotto numbers in such a manner. We used age as a proxy for experience, and education level as a proxy for the capability of understanding parimutuel games features. In addition, since our participants live in Belgium, which is a country with two different official languages, namely Dutch in Flanders and French in Wallonia, we added a control variable for language to take this specificity of our study context into account.¹⁰ Finally, following the discussion about the difference between regular and occasional or non-players (see Section 4.3), we added a dummy variable *Regular_Player* set to zero for participants who declared having never or rarely played lotto during the last year and one otherwise.

Table 3 provides the results of estimating binary logit regression models using two different dependent variables. In Model 1, the dependent variable is *Birthday_OR*, which is set to 1 when *Birthday_Numbers* or *Birthday_Others* is strictly greater than 4 (i.e., the midpoint of the scale). Hence, Model 1 examines the probability of declaring a tendency to select birthday numbers when playing lotto games (stated preference). In Model 2, the dependent variable is *Play_Date*, which is set to 1 when participants selected their birthdate day and/or month on the first lotto ticket. In contrast to Model 1, Model 2 relies on a direct measure of the birthday-number effect (revealed preference).

¹⁰ Past research has already reported some differences between Dutch-speaking and French-speaking people in Belgium when it comes to economic decisions and risk-taking (e.g., D'Hondt et al., 2021).

Results reported in Table 3 show that gender, which is characterized with a dummy variable (*Dummy_Female*) set to 1 for women, is significant in both models. This means that, on average, women are both more likely to declare their tendency to play birthday numbers and more likely to walk their talk (i.e., select their birthday numbers on the first lotto ticket). Looking at the odds ratio, the probability of declaring a tendency to select birthday numbers when playing lotto is multiplied by a factor of 2.144 for women, indicating that, all else being equal, the probability of stated preference for birthday numbers more than doubles for women, compared to men. In other words, women are twice more likely than men to declare a tendency to select birthday numbers when playing lotto. Regarding the probability of participants selecting their birthdate day and/or month on the first lotto ticket, the odds ratio is equal to 1.647, meaning that this probability is on average higher by about 65% for women, compared to men. Hence, both results provide strong support for a gender gap in the birthday-number effect, while controlling for other individual characteristics. Therefore, we find support for H2.

Table 3 reveals that some of the control variables exhibit significant coefficient estimates in Models 1 and 2, meaning that these variables help explain both stated and revealed preference for birthday numbers among participants. First, *Self_esteem* exhibits positive and significant coefficient estimates (at the 10% level, with corresponding odds ratios around 1.2), which reveals a positive relationship between perceived self-esteem and the birthday-number effect, all else being equal. This finding, which supports H3, is consistent with Kitayama and Karasawa (1997) and Koole et al. (2001) who consider the birthday-number effect as a valid indicator of self-esteem, alongside with the name-letter effect. Age is also a significant determinant in Models 1 and 2, although with opposite signs. It displays a negative relationship with the probability of stated preference for birthday numbers (at the 5% level) but a positive relationship with the probability of revealed preference for birthday

numbers (at the 10% level), all else being equal. However, in both cases, the odds ratio is close to one, meaning that age is a rather weak driver of the birthday-number effect.

Focusing on the other control variables exhibiting significant coefficient estimates only in Model 1, we gain further insights into the probability of stated preference for birthday numbers. First, *Superstition*, which captures the extent to which participants are subject to superstitious beliefs, displays a positive coefficient estimate (at the 1% level), with a corresponding odds ratio of 1.293 (meaning that, all else being equal, when superstitious beliefs increase by one unit on the seven-point Likert scale, the probability of stated preference for birthday numbers increases by 29.3%). This reveals a strong positive relationship between superstitious beliefs and the probability of stated preference for birthday numbers, which supports H4. Second, as expected, the level of education is negatively related to the probability of stated preference for birthday numbers. Specifically, when participants hold a university degree (or equivalent), the probability of declaring a tendency to play birthday numbers decreases by approximately 33% (odds ratio=0.668), all else being equal. Third, being in a relationship as a couple is positively associated with the probability of stated preference for birthday numbers. This dummy variable (set to 1 for participants who are not single, divorced, separated, or widow/widower) displays a positive coefficient estimate (at the 5% level) and an odds ratio of 1.616. When people are in a couple, the probability of declaring a tendency to play birthday numbers increases by about 61%, compared to people who are not in couple, all else being equal. In contrast, the variable *Children* is not significant for the probability of stated preference for birthday numbers. However, we should point out that the information captured by this variable might be noisy because the corresponding question in the survey referred to underage children only (i.e., kids less than 18 years old). It seems reasonable to assume that at least some participants are likely to have children older than 18 since the average participant is approximately 45 years old. The *Regular_Player*

dummy variable, though marginally significant (at the 10% level), has an odds ratio equal to 1.485. This result reveals that the probability of stated preference for birthday numbers is 48% higher for regular players than for occasional or non-players, which is consistent with the findings of Table A2 (in appendix). Finally, the dummy variable set to 1 for French-speaking participants (*Dummy_French*) has a positive coefficient estimate (at the 1% level), revealing that the probability of stated preference for birthday numbers is 205% higher for them (odds ratio=2.054) than for Dutch-speaking participants.

Table 3
Multivariate analysis of the birthday-number effect

Variables	Model 1 Birthday_OR		Model 2 Play_Date	
	Estimates	Odds ratio	Estimates	Odds ratio
<i>Intercept</i>	-1.3039		-1.2541*	
<i>Age</i>	-0.0171**	0.983	0.0120*	1.012
<i>Dummy_Female</i>	0.7629***	2.144	0.4988**	1.647
<i>Dummy_French</i>	0.7198***	2.054	0.2435	1.276
<i>In couple</i>	0.4802**	1.616	0.1322	1.141
<i>Children</i>	0.0374	1.038	0.0346	1.035
<i>High_education</i>	-0.4037**	0.668	-0.1957	0.822
<i>Self_esteem</i>	0.1957*	1.216	0.1855*	1.204
<i>Superstition</i>	0.2573***	1.293	-0.0134	0.987
<i>Regular_Player</i>	0.3957*	1.485	0.1876	1.206
<i>R</i> ²	0.1091		0.0346	
<i>N</i>	498		498	

Note: This table reports the results of estimating two binary logit regression models using the entire sample. In Model 1, the dependent variable is *Birthday_OR*, which is an indirect measure of the birthday-number effect set to 1 when *Birthday_Numbers* or *Birthday_Others* is strictly greater than 4. In Model 2, the dependent variable is *Play_Date*, which is a direct measure of the birthday-number effect set to 1 when participants selected their birthdate day and/or month on the first lotto ticket. *Dummy_Female* is a dummy variable for gender, which is set to 1 for women. *Dummy_French* is a dummy variable set to 1 for French-speaking participants. *In_couple* is another dummy variable set to 1 for participants who are not single, divorced, separated, or widow/widower. *Children* refers to the number of underage children reported by participants. *High_education* is a dummy variable when participants hold a university degree (or equivalent). *Self_esteem* is measured with Rosenberg's (1965) Self-Esteem Scale. *Superstition* refers to the superstitious personality trait, measured with the four items from Carlson et al. (2009). *Regular_Player* is a dummy variable set to 1 for regular players. *N* gives the sample size, i.e., the number of participants. Significance levels: *** $p < .01$, ** $p < .05$, * $p < .10$.

5.1 Robustness check

The findings reported in Table 3 might be somewhat biased because our sample is not perfectly balanced between men and women (238 women for 260 men). In addition, the distribution of birthdate days is not homogeneous among men and women (see Table 1), which might also affect our results. We therefore matched women and men using their birthdate days. We applied a matching without replacement, meaning that a given male participant cannot be matched to several female participants. We ended up with an after-match sample of 398 participants, which is perfectly balanced and homogeneous regarding the distribution of birthdate days.

Table 4 provides the results of estimating the same binary logit regression models used above. It appears that our matching reinforces the results in both models, without affecting the relationships with most of the control variables. For the probability of stated preference for birthday numbers (Model 1), the gender gap is strengthened since the odds ratio for gender (*Dummy_Female*) is higher in Table 4 than in Table 3 (i.e., 2.510 versus 2.144). Consistent evidence is found for the probability of revealed preference for birthday numbers (see Model 2), with the odds ratio for gender being equal to 1.878 in Table 4, instead of 1.647 in Table 3.

Table 4
Multivariate analysis of the birthday-number effect in the after- match sample

Variables	Model 1 Birthday_OR		Model 2 Play_Date	
	Estimate	Odds ratio	Estimate	Odds ratio
<i>Intercept</i>	-1.5365*		-1.6337**	
<i>Age</i>	-0.0199**	0.980	0.0114	1.011
<i>Dummy_Female</i>	0.9202***	2.510	0.6301***	1.878
<i>Dummy_French</i>	0.8783***	2.407	0.3096	1.363
<i>In_couple</i>	0.6194**	1.858	0.1713	1.187
<i>Children</i>	0.1948	1.215	0.0330	1.034
<i>High_education</i>	-0.4014*	0.669	-0.2385	0.788
<i>Self_esteem</i>	0.2054*	1.228	0.2569**	1.293
<i>Superstition</i>	0.1972**	1.218	-0.0450	0.956
<i>Regular_Player</i>	0.4253*	1.530	0.2930	1.340
<i>R</i> ²	0.1329		0.0513	
<i>N</i>	398		398	

Note: This table reports the results of estimating two binary logit regression models using the after-match sample. In Model 1, the dependent variable is *Birthday_OR*, which is an indirect measure of the birthday-number effect set to 1 when *Birthday_Numbers* or *Birthday_Others* is strictly greater than 4. In Model 2, the dependent variable is *Play_Date*, which is a direct measure of the birthday-number effect set to 1 when participants selected their birthdate day and/or month on the first lotto ticket. *Dummy_Female* is a dummy variable for gender, which is set to 1 for women. *Dummy_French* is a dummy variable set to 1 for French-speaking participants. *In_couple* is another dummy variable set to 1 for participants who are not single, divorced, separated, or widow/widower. *Children* refers to the number of underage children reported by participants. *High_education* is a dummy variable when participants hold a university degree (or equivalent). *Self_esteem* is measured with Rosenberg's (1965) Self-Esteem Scale. *Superstition* refers to the superstitious personality trait, measured with the four items from Carlson et al. (2009). *Regular_Player* is a dummy variable set to 1 for regular players. *N* gives the sample size, i.e., the number of participants. Significance levels: *** $p < .01$, ** $p < .05$, * $p < .10$.

5.2 On the persistence of the birthday-number effect in sequential choices

So far, we have been using the direct measures of the birthday-number effect relying on the first lotto tickets filled in by participants. To address whether the birthday-number effect is persistent in sequential choices (see H5), we are now interested in the comparison of the observed values on the two successive tickets filled in by participants. As mentioned earlier, in the online survey, they are requested to fill in the second lotto ticket immediately after the first one (i.e., participants completed the two lotto tickets in a row).

Table 5 shows the results of univariate comparisons between the first and second lotto tickets that participants completed for the following three variables *Mean_Draw*, *Less30*, and *Birthday_Day*. In addition, we added the results for two other variables aiming at measuring the time-to-completion. Specifically, the first variable is *Lotto_Duration*, which measured the time (in seconds) taken by the participant to fill in the lotto ticket. The second variable is *Click_Count*, which counted the excess number of clicks between the first and the last click on the ticket, that is, the difference between the actual number of clicks minus the minimum number of clicks required to fill in the ticket (equal to 6 for the Belgian lotto game and in our setting). This variable is intended to be a proxy measure for capturing the participant's excess of deliberation through changes of mind and hesitations when selecting numbers.

Table 5
Test of the persistence of the birthday-number effect

Variables	Ticket 1	Ticket 2	t-stat	p-value
<i>N</i>	498	498		
<i>Mean_Draw</i>	19.49	20.92	-4.46	<.0001***
<i>Less30</i>	4.66	4.40	4.18	<.0001***
<i>Birthday_Day</i>	0.59	0.14	16.60	<.0001***
<i>Lotto_Duration</i>	31.88	19.76	8.00	<.0001***
<i>Click_Count</i>	1.98	1.27	3.73	0.0002***

Note: This table reports cross-sectional means among the first and the second lotto ticket as well as the results of statistical tests comparing these means. *Mean_Draw* is the average of the numbers selected by participants on the lotto tickets. *Less30* counts participants' selected numbers lower or equal to 30 on the lotto tickets. *Birthday_Day* is a dummy variable set to 1 when participants selected their birthdate day on the lotto tickets. *Lotto_Duration* is a measure of the time (in seconds) used by participants to fill in the lotto tickets. *Click_Count* refers to the excess number of clicks between the first and the last click on the ticket, that is, the difference between the actual number of clicks minus the minimum number of clicks required to fill in the lotto ticket (6 in our setting, like in the Belgian lotto game). *N* gives the number of participants. Significance levels: *** $p < .01$, ** $p < .05$, * $p < .10$. For *Lotto_Duration*, one participant who took 3,176 seconds to fill in the second ticket was disregarded (*N* is then equal to 497 for that variable).

Before interpreting the results of Table 5, it is worth recalling that the theoretical mean of the numbers selected on each lotto ticket under random choice is $\mu = \frac{1+45}{2} = 23$. For the first lotto tickets, the sample average for *Mean_Draw* is 19.49, which is significantly lower than 23 ($t(497) = -15.33, p < .0001$). For the second tickets, the average of *Mean_Draw* is somewhat higher (20.92), although it remains significantly below 23 ($t(497) = -9.56, p < .0001$). However, there is a significant increase for *Mean_Draw* between tickets 1 and 2 ($t(497) = -4.46, p < .0001$). Hence, while conscious selection was still working, participants selected on average higher numbers on their second lotto ticket, compared to the first one.

Regarding the variable *Less30*, we observe the same kind of result. The sample average is equal to 4.66 for ticket 1 and 4.40 for ticket 2. As players selected 6 numbers between 1 and 45, there should be on average 4 numbers below 30 (out of 6) if choices were made at random. Unsurprisingly, actual number choices reveal that the average of *Less30* is significantly larger than 4 on both tickets ($Less30 = 4.66, t(497) = 15.09, p < .0001$ for ticket 1 and $Less30 = 4.40$

$t(497) = 8.92, p < .0001$ for ticket 2). However, the mean of *Less30* is significantly lower for the second ticket, compared to the first ticket ($t(497) = 4.18, p < .0001$).

Of particular interest, Table 5 indicates the complete disappearance of the birthday-number effect on the second lotto ticket, which supports H5. While almost 60% of participants selected their birthdate day on the first ticket, only 14% of them replicated this choice when filling in the second ticket. Since this proportion does not significantly differ from the theoretical proportion under random choice (13.33%), this finding does not provide support for the presence of the birthday-number effect on the second lotto ticket completed by participants.

Table 5 also shows that, on average, participants took 31.88 seconds to fill in their first ticket while they only needed 19.76 seconds to select numbers on their second ticket.¹¹ The difference between these two durations is highly significant ($t(496) = 8.00, p < .0001$). Similarly, participants made more changes and hesitated more when selecting numbers on the first ticket, as shown by the results for the variable *Click_Count*. Its average value is equal to 1.98 for the first ticket, against 1.27 for the second ticket, which reveals a significant decrease in the number of clicks ($t(497) = 3.73, p = 0.0002$). In terms of both number of seconds and number of clicks, participants completed the second lotto ticket more quickly.

Consistent with the aforementioned disappearance of the birthday-number effect, the gender gap also vanished when considering the second lotto tickets. Table 6 provides the results of univariate tests comparing men and women using the second lotto tickets. When comparing the results of Table 6 with those reported in Table 2 using the first lotto tickets, the only difference that still emerges between men and women in Table 6 refers to the variable *Less30*. The sign of this difference in Table 6 is the opposite of the one reported in Table 2. Hence, when filling in the second ticket, men selected more numbers lower than 30, compared to women, which is just the contrary of what was found when considering the first lotto tickets.

¹¹ For this specific test, one participant was considered as missing because he took almost one hour to fill in the second ticket. Therefore, the test is performed on 497 participants.

Table 6
Univariate analysis of gender gap in the birthday-number effect
(second lotto ticket)

Variable	Men	Women	t-value	p-value
<i>N</i>	260	238		
<i>Mean_Draw</i>	20.72	21.14	-0.96	0.3355
<i>Less30</i>	4.47	4.32	1.68	0.0929*
<i>Birthday_Day</i>	0.13	0.15	-0.53	0.5965
<i>Birthday_Month</i>	0.16	0.17	-0.31	0.7546

Note: This table reports cross-sectional means and proportions among men and women as well as the results of statistical tests comparing these means and proportions. *Mean_Draw* is the average of the numbers selected by participants on their second lotto ticket. *Less30* counts participants' selected numbers lower or equal to 30 on their second lotto ticket. *Birthday_Day* and *Birthday_Month* are dummy variables set to 1 when participants selected either their birthdate day or month on their second lotto ticket. *N* gives the number of participants. Significance levels: *** $p < .01$, ** $p < .05$, * $p < .10$.

To further check in a multivariate setting whether the birthday-number effect disappeared when people filled in their second lotto ticket, we estimated another binary logit regression model, wherein the dependent variable *Play_Date* is set to 1 when participants selected their birthdate day and/or month on their second lotto ticket. The results are reported in Table 7, when estimating this model using either the entire sample or the after-match sample. In both cases, the coefficient estimates for gender (see *Dummy_Female*) are not significant at any conventional levels, confirming that the birthday-number effect vanished when participants, regardless of their gender, selected numbers on the second lotto tickets.

Table 7
Multivariate analysis of the birthday-number effect using the second lotto tickets

Variables	Play_Date (entire sample)		Play_Date (after-match sample)	
	Estimate	Odds Ratio	Estimate	Odds Ratio
<i>Intercept</i>	-0.7215		-1.1939	
<i>Age</i>	-0.0137*	0.986	-0.0189**	0.981
<i>Dummy_Female</i>	0.2365	1.267	0.3273	1.387
<i>Dummy_French</i>	0.3783*	1.460	0.4345*	1.544
<i>In_couple</i>	0.5012**	1.651	0.5032**	1.654
<i>Children</i>	0.0525	1.054	0.0983	1.103
<i>High_education</i>	-0.1594	0.853	-0.1886	0.828
<i>Self_esteem</i>	-0.0043	0.996	0.1109	1.117
<i>Superstition</i>	-0.0834	0.920	-0.0803	0.923
<i>Regular_Player</i>	0.0470	1.048	0.0734	1.076
<i>R</i> ²	0.0299		0.0433	
<i>N</i>	498		398	

Note: This table reports the results of estimating a binary logit regression model wherein the dependent variable is *Play_Date*, which is a direct measure of the birthday-number effect set to 1 when participants selected their birthdate day and/or month on their second lotto ticket. This model was estimated using the entire sample (results on the left-hand side) and using the after-match sample (results on the right-hand side). *Dummy_Female* is a dummy variable for gender, which is set to 1 for women. *Dummy_French* is a dummy variable set to 1 for French-speaking participants. *In_couple* is another dummy variable set to 1 for participants who are not single, divorced, separated, or widow/widower. *Children* refers to the number of underage children reported by participants. *High_education* is a dummy variable when participants hold a university degree (or equivalent). *Self_esteem* is measured with Rosenberg's (1965) Self-Esteem Scale. *Superstition* refers to the superstitious personality trait, measured with the four items from Carlson et al. (2009). *Regular_Player* is a dummy variable set to 1 for regular players. *N* gives the sample size, i.e., the number of participants. Significance levels: *** $p < .01$, ** $p < .05$, * $p < .10$.

6. Conclusion

This paper adds to the literature on conscious selection by making several contributions. First, we provide evidence of the birthday-number effect among Belgian lotto players, using both indirect and direct measures based on survey data. Most importantly, we reveal the existence of a gender gap in this effect when participants fill in the first lotto ticket. All our findings show that, on average, women are not only more likely to declare their tendency to play birthday numbers, but they are also more likely to select their own birthday numbers on the first lotto ticket. Our multivariate results indicate that this gender gap is meaningful, even when

controlling for individual characteristics like socio-demographic indicators (e.g., age, education, having children or being single) and relevant personality traits like self-esteem and superstitious beliefs.

Next, we show the non-persistence of the birthday-number effect, and consecutively of the gender gap, when people are asked to fill in a second lotto ticket (directly after the first one). Interestingly, the behavior of participants changes when they fill in this second ticket, despite the absence of any differences in the instructions provided to them. In particular, we find a strong decrease in the time-to-completion, though players continue to select smaller numbers on average, compared to a random choice. Taken together, these findings suggest that the birthday-number effect is the result of a *deliberative* process since participants betting on their birthday numbers take significantly more time to select numbers. Contrary to Koole et al. (2001), who report that the positive bias for birthday numbers is inhibited when participants are induced to respond in a deliberative manner, in our lotto game setting the strength of the birthday-number effect is positively related to *deliberative* number choices, not to intuitive and automatic number selection. Hence, we argue that the underlying mental process behind the birthday-number effect in lottery gambling is *more reflective* than intuitive because the birthday-number effect disappears when people select numbers “without thinking” and significantly faster on the second lotto ticket.

Our findings, as ecologically valid as they may be, are exclusively based on survey data and might therefore not fully reflect the dynamics and complexity of the decisions lotto players make in reality. We show that the birthday-number effect is a strong driver of conscious selection, and consequently a suboptimal way to choose numbers in a parimutuel game. The striking result that 55.77% of men and 63.45% of women play their birthdate day when asked to fill in a first lotto ticket calls, however, for further investigations relying on field data, using,

for example, the registration information of online gamblers.¹² Moreover, addressing the birthday-number effect using field data would allow to study whether people filling in several tickets notch their numbers faster on the second or third ticket, compared to the first one, as observed in our survey data. Response times have become recently a topic explored by economists (Allos-Ferret et al., 2021; Clithero, 2018; Liu, 2023). To the best of our knowledge, they are not yet used in gambling studies despite the information they reveal. The second limitation of our work is that we do not study the process that leads people to bet on birthdate numbers. A qualitative study based on interviews would be useful to understand whether playing on birthdates is based on the identification of those dates as “lucky” numbers or if these numbers are simply easier to remember, as suggested by Baker and McHale (2009, 2011).

This paper paves the way to future research since our methodology calls for disentangling the preference for small numbers into several components, the most obvious being the birthday-number effect. The other components might cover superstitious beliefs (lucky numbers are in general “small” numbers in the lotto context) or geometric/aesthetic considerations (e.g., selecting numbers to form special shapes on the ticket like diagonals, complete rows, etc.) when filling in the ticket form (Falk et al. 2009; Wang et al., 2016).

¹² For example, people who register on the Belgian lottery website are required to provide their birthdate (<https://www.loterie-nationale.be/enregistrement/devenir-membre>).

References

- Auriat, N. (1993). "My wife knows best" A comparison of event dating accuracy between the wife, the husband, the couple, and the Belgium population register. *Public Opinion Quarterly*, 57 (2), 165-190.
- Baker, R.D. and McHale, I.G. (2009). Modelling the probability distribution of prize winnings in the UK National Lottery: consequences of conscious selection. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172, 813-834.
- Alos-Ferrer C., Fehr, E. and Netzer, N. (2021). Time will Tell: Recovering Preferences When Choices are Noisy. *Journal of Political Economy*, 129, 6, 1828-1877.
- Baker, R. and McHale, I.G. (2011). Investigating the behavioural characteristics of lottery players by using a combination preference model for conscious selection. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174, 1071-1086.
- Bao, W. Wang, Y., Yu, T., Zhou, J., and Luo, J. (2022). Women rely on "gut feeling"? The neural pattern of gender difference in non-mathematic intuition, *Personality and Individual Differences*, 196, 111720.
- Carlson, B. D., Mowen, J. C., and Fang, X. (2009). Trait Superstition and Consumer Behavior: Re-Conceptualization, Measurement, and Initial Investigations. *Psychology & Marketing*, 26(8), 689-713.
- Clithero, J. A. (2018). Response Times in Economics: Looking through the Lens of Sequential Sampling Methods. *Journal of Economic Psychology*, 69, 61-86.
- Clotfelter, C. T., & Cook, P. J. (1993). Notes: The gambler's fallacy in lottery play. *Management Science*, 39, 1521-1525.
- Cook, P. J., & Clotfelter, C. T. (1993). The peculiar scale economies of Lotto. *The American Economic Review*, 83, 634-643.
- Coulter, K.S. and Grewal, D. (2014). Name-Letters and Birthday-Numbers: Implicit Egotism Effects in Pricing, *Journal of Marketing*, 78(3), 102-120.
- Craighead, C. W., Ketchen, D. J., Dunn, K. S., and Hult, G. T. M. (2011). Addressing Common Method Variance: Guidelines for Survey Research on Information Technology, Operations, and Supply Chain Management. *IEEE Transactions on Engineering Management*, 58(3), 578-588.
- D'Hondt, C., McGowan, R., and Roger, P. (2021), Trading leveraged Exchange-Traded products is hazardous to your wealth, *The Quarterly Review of Economics and Finance*, 80, 287-302.
- Falk, R., Falk, R., & Ayton, P. (2009). Subjective patterns or randomness and choice: Some consequences of collective responses. *Journal of Experimental Psychology: Human Perception and Performance*, 35(1), 203-224.

Farrell, L., R. Hartley, G. Lanot, and I. Walker (2000). The Demand for Lotto: The Role of Conscious Selection, *Journal of Business and Economic Statistics*, 18(2), 228–241.

Fiebert, M. Tilmont, L. and Warren, C.R. (2013). It's Your Birthday!: Greetings as a Function of Gender and Relationship Status on Facebook, *International Review of Social Sciences and Humanities*, 4(2), 206-208.

Finkelstein, M. (1995). Estimating the frequency distribution of the numbers bet on California lottery, *Applied Mathematics and Computation*, 69, 195-207.

He, J., Liu, H., Sing, T. F., Song, C., Wong, W.-K.(2020). Superstition, conspicuous spending, and housing market: Evidence from Singapore. *Management Science*, 66, 783– 804.

Hirshleifer, D., Jian, M., Zhang, H.(2018). Superstition and financial decision making. *Management Science*, 64, 235–252.

Husemann-Kopetzky, M. and Köcher, S. (2016). Price Endings that Matter: A Conceptual Replication of Implicit Egotism Effects in Pricing. *Journal of Marketing Behavior*, 2, 313–324.

Jones, J. T., Pelham, B. W. Mirenberg, M. C. and Hetts, J. J. (2002) Name Letter Preferences Are Not Merely Mere Exposure: Implicit Egotism as Self Regulation, *Journal of Experimental Social Psychology*, 20(2), 170-177

Kitayama, S., Karasawa, M. (1997). Implicit self-esteem in Japan: Name letters and birthday numbers, *Personality and Social Psychology Bulletin*, 23(7), 736-742.

Koole, S. L., Dijksterhuis, A., and van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology*, 80(4), 669–685.

Liu, S., and Netzer, N. (2023). Happy Times: Measuring Happiness Using Response Times. *American Economic Review*, 113 (12), 3289-3322.

Netemeyer, R. G., Bearden, W. O., and Sharma, S. (2003). *Scaling Procedures: Issues and Applications*. Thousand Oaks, CA: Sage Publications.

Nuttin, J.M., Jr. (1985). Narcissism beyond gestalt and awareness: The name letter effect. *European Journal of Social Psychology*, 15, 353-361.

Pelham, B. W., Koole, S. L., Hardin, C. D., Hetts, J. J., Seah, E., and DeHart, T. (2005). Gender moderates the relation between implicit and explicit self-esteem. *Journal of Experimental Social Psychology*, 41, 84-89.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., and Podsakoff, N. P. (2003). Common Method Bias in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, 88(5), 879-903.

Polin, B.I., and Aharon, I. (2021). Patterns in manually selected numbers in the Israeli lottery. *Judgment and Decision Making*, 16(4), 1039-1059.

Roger, P., and Broihanne, M.-H. (2007). Efficiency of betting markets and rationality of players: Evidence from the French 6/49 Lotto. *Journal of Applied Statistics*, 34, 645–662.

Roger, P., D'Hondt, C., Plotkina, D., and Hoffmann, A. (2023),. Number 19: Another Victim of the COVID-19 Pandemic? *Journal of Gambling Studies*, 39(3), 1417–1450.

Rogers, P. (2008). The psychology of lottery gambling, *Journal of Gambling Studies*, 14(2), 111-134.

Rosenberg, M. (1965). *Rosenberg Self-Esteem Scale*. Princeton, NJ: Princeton University Press.

Simmering, M. J., Fuller, C. M., Richardson, H. A., Ocal, Y., and Atinc, G. M. (2015). Marker Variable Choice, Reporting, and Interpretation in the Detection of Common Method Variance: A Review and Demonstration. *Organizational Research Methods*, 18(3), 473-511.

Simon, J. (1999) An analysis of the distribution of combinations chosen by UK National Lottery players, *Journal of Risk and Uncertainty*, 17(3), 243–276.

Stetzka, R. M. and Winter, S. (2021) How rational is gambling? *Journal of Economic Surveys*, 1– 57.

Turner, N. E. (2010). Lottery ticket preferences as indicated by the variation in the number of winners. *Journal of Gambling Studies*, 26, 421–439.

Wang, T. V., van Loon, R. J. P., van den Assem, M. J., and van Dolder, D. (2016). Number preferences in lotteries. *Judgment and Decision Making*, 11, 243–259.

Appendix

Table A1: Measurement description and reliability of self-esteem and superstitious beliefs

Scale/Variable	Items	Mean (SD)	Factor loading	α	CR	AVE
Explicit self-esteem	<i>To which extent do you agree with the following statements. From 1 = "Completely disagree" to 7 = "Completely agree."</i>					
(Rosenberg 1965)	Total scale	4.87 (1.04)		.886	.790	.327
	1. On the whole, I am satisfied with myself.	5.18 (1.24)	.349			
	2. At times I think I am no good at all. (R)	4.40 (1.66)	.770			
	3. I feel that I have a number of good qualities.	5.32 (1.09)	.100			
	4. I am able to do things as well as most other people.	5.18 (1.25)	.150			
	5. I feel I do not have much to be proud of. (R)	5.02 (1.59)	.605			
	6. I certainly feel useless at times. (R)	4.37 (1.80)	.811			
	7. I feel that I am a person of worth, at least on an equal plane with others.	5.20 (1.26)	.183			
	8. I wish I could have more respect for myself. (R)	3.94 (1.68)	.787			
	9. All in all, I am inclined to feel that I am a failure. (R)	5.10 (1.74)	.756			
	10. I take a positive attitude toward myself.	5.01 (1.38)	.530			
Superstitious beliefs	Total scale	3.19 (1.61)		.795	.865	.623
(Carlson et al., 2009)	1. I sometimes perform little rituals to bring good luck to myself.	3.35 (1.91)	.879			
	2. I do not want to lose things that bring me good luck.	5.34 (1.60)	.544			
	3. I have to admit that sometimes I act like I am superstitious.	3.37 (1.94)	.838			
	4. I sometimes bring a lucky charm with me.	3.44 (1.93)	.851			

Notes. SD = Standard Deviation. α = Cronbach's alpha. CR = Composite Reliability. AVE = Average Variance Extracted; R = reversed items.

Table A2
Univariate analysis of gender gap in the birthday-number effect
for regular players and occasional or non-players
(first lotto tickets)

Variable	Men	Women	t-value	p-value
PANEL A: Occasional or non-players				
<i>N</i>	73	112		
<i>Mean_Draw</i>	20.48	19.02	1.93	0.0553*
<i>Less30</i>	4.47	4.77	-2.08	0.0386**
<i>Birthday_Numbers</i>	3.89	5.08	-3.89	0.0001***
<i>Birthday_Others</i>	3.66	5.05	-4.46	<.0001***
<i>Birthday_Day</i>	0.52	0.63	-1.53	0.1269
<i>Birthday_Month</i>	0.33	0.39	-0.88	0.3796
PANEL B: Regular players				
<i>N</i>	187	126		
<i>Mean_Draw</i>	19.70	19.03	1.13	0.2606
<i>Less30</i>	4.62	4.75	-1.11	0.2692
<i>Birthday_Numbers</i>	4.79	5.23	-2.03	0.0434**
<i>Birthday_Others</i>	4.70	5.28	-2.66	0.0081***
<i>Birthday_Day</i>	0.57	0.63	-1.11	0.2685
<i>Birthday_Month</i>	0.40	0.44	-0.62	0.5342

Note: This table reports cross-sectional means and proportions among men and women as well as the results of statistical tests comparing these means and proportions. *Mean_Draw* is the average of the numbers selected by participants on their first lotto ticket. *Less30* counts participants' selected numbers lower or equal to 30 on their first lotto ticket. *Birthday_Numbers* refers to the level selected on a seven-point Likert scale assessing participants' self-declared tendency to select their own birthdate numbers when playing lotto. *Birthday_Others* refers to the level selected on a seven-point Likert scale assessing participants' self-declared tendency to select birthday numbers of other people when playing lotto. *Birthday_Day* and *Birthday_Month* are dummy variables set to 1 when participants selected either their birthdate day or month on their first lotto ticket. Regular players (313 participants) declared playing at least once a month (up to several times per week) while occasional players or non-players declared playing less than once a month or never (185 participants). *N* gives the number of participants. Significance levels: *** $p < .01$, ** $p < .05$, * $p < .10$.