



dECON

Facultad de Ciencias Sociales
UNIVERSIDAD DE LA REPÚBLICA

Documentos de Trabajo

Smokers' Rational Lexicographic Preferences for Cigarette Package Warnings: A Discrete Choice Experiment with Eye Tracking

Jeffrey E. Harris, Mariana Gerstenblüth y Patricia Triunfo

Documento No. 02/18
Setiembre 2018

ISSN 0797-7484

Smokers' Rational Lexicographic Preferences for Cigarette Package Warnings: A Discrete Choice Experiment with Eye Tracking

JEFFREY E. HARRIS, MARIANA GERSTENBLÜTH AND PATRICIA TRIUNFO*

*Harris: Department of Economics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, E52-422, Cambridge MA 02139 and NBER (email: jeffrey@mit.edu); Gerstenblüth: Departamento de Economía, Facultad de Ciencias Sociales, Universidad de la República (UDELAR), Constituyente 1502, 11200 Montevideo, Uruguay (email: mariana.gerstenbluth@cienciassociales.edu.uy); Triunfo: Departamento de Economía, Facultad de Ciencias Sociales, UDELAR, Constituyente 1502, 11200 Montevideo, Uruguay (email: patricia.triunfo@cienciassociales.edu.uy). We thank Gastón Ares (Instituto Polo Tecnológico de Pando, Facultad de Química, UDELAR) and Leandro Machin (Centro de Investigación Básica en Psicología, Facultad de Psicología, UDELAR) for their assistance in conducting the experiment and their insightful comments. We thank Alejandra Arrúa (Facultad de Psicología, UDELAR) and Sebastián Suárez (Facultad de Arquitectura, Diseño y Urbanismo, UDELAR) for the package illustrations. We gratefully acknowledge the financial support of the Unión Contra la Tuberculosis y Enfermedades Respiratorias through an unrestricted grant to the Tobacco Control Program of the Ministry of Public Health of Uruguay. The authors have no competing interests to declare.

ABSTRACT

We asked 97 cigarette smokers to make a series of 12 binary choices between experimental cigarette packages with varying warnings and background colors. Each smoker had to decide which of the two packages contained cigarettes less risky for his health. We tested whether the smokers, confronted with warnings that were repugnant and threatening to many of them, could still make choices that adhered to the standard axioms of rational choice. We supplemented our observations on smokers' choices with data on their eye movements. We find that participants universally made choices consistent with a complete, transitive preference ordering. We find little evidence of inconsistent choices violating the weak axiom of revealed preference. In a majority of smokers, we find strong evidence of the use of a lexicographic decision rule to assess the riskiness of a cigarette package. These smokers first ranked the two packages solely on the basis of their warnings. Only when the two packages had the same warning did they rank the packages on the basis of their color. The data on eye tracking strongly confirmed the lexicographic nature of the underlying decision rule. Our study represents an entirely different angle of inquiry into the question of rational addiction.

Key words: addiction, cigarettes, smoking, health warnings, rationality, discrete choice experiment, eye tracking, transitivity, additive utility, lexicographic preferences, context-dependent preferences, response time, drift diffusion model, Schelling-Thaler-Shefrin dual-self model.

JEL: D12, D83, D87, D91, I12, M31.

RESUMEN

Realizamos un experimento con 97 fumadores de cigarrillos, los cuales se enfrentaron a 12 opciones binarias de paquetes de cigarrillos con diferentes advertencias y colores de fondo. En cada caso, cada fumador tuvo que decidir cuál de los dos paquetes consideraba contenía el producto menos riesgoso para su salud. Analizamos si los fumadores, al enfrentarse a advertencias que eran repugnantes y amenazantes, aún podían tomar decisiones consistentes con los axiomas de la elección racional. Complementamos nuestras observaciones sobre las elecciones de los fumadores con datos de sus movimientos oculares obtenidos a través de la técnica de *eye tracking*. Encontramos que los participantes tomaron decisiones universalmente consistentes con un ordenamiento de preferencias completo y transitivo, así como poca evidencia de elecciones inconsistentes que violen el axioma débil de las preferencias reveladas. En la mayoría de los fumadores, encontramos pruebas contundentes del uso de una regla de decisión lexicográfica para evaluar el riesgo de un paquete de cigarrillos. Estos fumadores primero clasificaron los dos paquetes únicamente en base a las advertencias, y sólo cuando los dos paquetes tuvieron la misma advertencia clasificaron los paquetes en función de su color. Los datos de *eye tracking* confirmaron la naturaleza lexicográfica de la decisión subyacente. Consideramos que nuestro estudio respresenta una perspectiva completamente diferente en el estudio de la adicción racional.

Palabras clave: adicción, cigarrillos, advertencias, racionalidad, experimentos de elección discreta, *eye tracking*, transitividad, utilidad aditiva, preferencias lexicográficas, preferencias dependientes del contexto, tiempo de respuesta, modelos *drift diffusion*, modelo de Schelling-Thaler-Shefrin.

In a landmark report released in May 1988, the U.S. Surgeon General concluded that the nicotine in cigarette smoke was an addictive drug. The report described cigarette smoking as a highly controlled, compulsive behavior driven by strong, often irresistible urges that could overcome repeated attempts to quit (U.S. Department of Health and Human Services 1988). In an article published just three months later, economists Gary Becker and Kevin Murphy inquired whether a cigarette smoker could still make rational decisions within the confines of his addiction (Becker and Murphy 1988).

Prior economic research on rational addiction has focused mostly on the smoker's ability to make farsighted rather than myopic utility-maximizing consumption choices (Gruber and Koszegi 2001). Here, we report on the results of a discrete choice experiment that looks at a different angle of rationality. We asked cigarette smokers to make a series of binary choices between experimental cigarette packages with varying warnings and background colors. With each successive choice, a participating smoker had to decide which of the two packages contained the cigarettes less risky for his health. We tested whether the participating smokers, confronted with experimental warnings that were undoubtedly repugnant and threatening to many of them, could still make choices that adhered to the classic axioms of rational choice. We supplemented our observations on smokers' choices with data on their eye movements, a technique that has attracted increasing interest among economists (Reutskaja et al. 2011, Lahey and Oxley 2016).

We find that participants in our experiment universally made choices consistent with a complete, transitive preference ordering. We find little evidence of inconsistent choices violating the weak axiom of revealed preference. What's more, in a majority of smokers, we find strong evidence of the use of a lexicographic decision rule to assess the riskiness of a cigarette package. These smokers first ranked the two packages solely on the basis of their warnings. Only when the two packages had the same warning did they rank the packages on the basis of their color. To be sure, our observations on the participants' choices were consistent with an additively separable utility function in which package warnings simply had far greater weight than package color. The additional data on eye tracking, however, strongly confirmed the lexicographic nature of the underlying decision rule. The smoker's eyes, in effect, provided a window into the neural sequences in his brain.

I. Experimental Design

A. Participants

We recruited a convenience sample of 98 self-reported adult current cigarette smokers aged 19–60 years from the students, faculty and staff of the Universidad de la República in Montevideo, Uruguay, to participate in a discrete choice experiment concerning their evaluation of the risks of different types of cigarettes.

Table 1 below shows the descriptive statistics for the 97 participants who completed the entire experiment. These study participants were predominantly female (nearly 60 percent). Most (68 percent) were under 30 years of age. Nearly half (46 percent) were students. Sixty percent smoked their first cigarette of the day within one hour of arising in the morning. Eighty-four percent reported smoking 20 or few cigarettes daily. Twenty-nine percent of the participants had made a serious attempt to quit smoking in the past year. Ninety-four percent believed that smoking caused lung cancer and 91% believed that it was a cause of heart disease.

All study participants reported normal or corrected-to-normal vision as well as full-color vision. Participants gave written informed consent and received a gift equivalent to US\$13 for participating in the study. The Ethics Committee of the School of Chemistry of the Universidad de la República, Uruguay, approved our study protocol.

B. Experimental Task

The experimental task consisted of a series of 12 predetermined choice sets shown consecutively on a computer monitor. Each choice set contained images of two cigarette packs that varied in design only along two specific dimensions: the warning, which consisted of an image and accompanying text occupying 80 percent of the face of the pack, and the background color. For each binary choice set shown on the computer screen, the participant was asked to click his mouse pointer on the cigarette pack that was “less risky for your health” (“menos riesgosa para su salud”). We adopted a forced-choice design. Participants were not permitted to opt out by indicating “none of these,” “not sure,” or “don’t know.” There was no time limit to make a choice. Our study design was similar to that of a previously reported discrete choice experiment on cigarette packages (Harris et al. 2017). In that study, however, we did not track eye movements.

TABLE 1—DESCRIPTIVE STATISTICS FOR 97 STUDY PARTICIPANTS*

Variable	Mean	SD
Female (%)	59.8	
Age (years)	28.2	8.8
Education attained:		
Some high school (%)	1.0	
Completed high school (%)	1.0	
Technical education (%)	6.2	
Some university (%)	75.3	
Graduated university (%)	8.3	
Graduate study (%)	8.3	
Cigarettes smoked per day:		
1–10 (%)	44.3	
11–20 (%)	40.2	
21–30 (%)	10.3	
More than 30 (%)	5.2	
Time to first cigarette on arising in a.m.:		
0–5 min. (%)	12.4	
6–30 min (%)	39.2	
31–60 min. (%)	18.6	
More than 60 min. (%)	40.2	
No response (%)	2.1	
Tried seriously to quit during past year (%)	28.9	
Believes that smoking causes lung cancer (%)	93.8	
Believes that smoking causes heart disease (%)	90.7	
Believes that smoking causes syphilis (%)	14.4	

*Includes 97 of 98 recruited subjects who completed the entire experiment. All participants were self-reported current cigarette smokers.

We randomized both the sequential order of the 12 choice sets and the right-left orientation of the two packs shown on the computer screen for each choice set. Specifically, among the 97 smokers who completed the entire experiment, 52 subjects were randomly assigned to Group I, while the remaining 47 subjects were randomly assigned to

Group II. Both groups were exposed to the same 12 choice sets in random order, but the right-left orientation of the two packs in each computer screen shown to Group II was the reverse of that shown to Group I.

Figure 1 below provides two examples of the 12 choice sets displayed to each participant on the computer monitor. For expositional purposes, we have labeled the twelve choice sets A through K. The figure specifically shows choice sets D and K displayed to the 52 participants randomized to Group I. Those in Group II, as noted above, saw the same packs, but with the right-left orientation reversed. The participants saw only what is shown within each rectangle, and not the labels above the rectangles.



FIGURE 1. CHOICE SETS D AND K SHOWN TO PARTICIPANTS IN GROUP I.

In choice set D, the warning on the left contained the image of a fetus held in a doctor's gloved hand, accompanied by the text, "Smoking during pregnancy harms the health of your baby." In the same choice set, the warning on the right contained the image of a tagged cadaver in a morgue, accompanied by the text, "Smoking causes heart attack." In choice set K, the warning on the left contained the image of a boot stamping out cigarettes, accompanied by the text, "Take the first step today. It's possible to quit smoking." The warning on the right contained the image of a mouth with an ulcer along the gum line, accompanied by the text, "Smoking causes bad breath, tooth loss and cancer of the mouth." The four warnings were selected from a public repository maintained by the Comisión Intergubernamental para el Control del Tabaco del MERCOSUR (CICT 2016), and had not appeared on any cigarettes marketed in Uruguay.

Each pack had one of three background colors: gray (RGB 122 136 129), light brown (RGB 167 115 51), or dark brown (RGB 74 65 42). Thus, in screen D, the pack at the left has a light brown background, while the pack at the right has a gray background. In screen K, the pack at the left has a dark brown background, while the pack at the right has a gray background. The dark brown background color, also referred to as Pantone 448C, is mandated by the Australian government on all packages of cigarettes sold in that country (Australian Government 2011). In a study of Australian smokers (Parr et al. 2011) that included the three background colors studied here, the dark brown was identified as implying the greatest harm.

Aside from the warning and the background color, all cigarette packs had the same design. The brand name, which did not correspond to any brand sold in Uruguay, was displayed in a uniform font, without a logo or other distinctive design element, at a fixed location at the bottom center of the pack. A skull-and-crossbones icon was placed at the bottom right, with the annotation, “Toxic Product.” The lateral edge contained the text warning, “This product contains nicotine, tar and carbon monoxide.” The overall design conformed to the requirements of plain packaging, currently in effect in Australia, France, the United Kingdom, New Zealand and Norway (Australian Government 2011, Moodie et al. 2018), and recently promulgated by the Uruguayan government.

Table 2 below shows the design parameters of the 12 different choice sets shown to the 52 participants randomized to Group I. The 12 choice sets were generated by the mix-and-match procedure (Johnson et al. 2007) implemented in the support.CEs package of R software (Aizaki 2012).

TABLE 2—CHOICE SETS SHOWN TO PARTICIPANTS IN GROUP I

Screen	Package on the Left	Package on the Right
A	Mouth, Dark Brown	Cadaver, Dark Brown
B	Fetus, Dark Brown	Fetus, Light Brown
C	Cadaver, Light Brown	Boot, Light Brown
D	Fetus, Light Brown	Cadaver, Gray
E	Mouth, Gray	Boot, Dark Brown
F	Boot, Gray	Fetus, Gray
G	Cadaver, Light Brown	Cadaver, Dark Brown
H	Fetus, Gray	Fetus, Dark Brown
I	Cadaver, Gray	Mouth, Dark Brown
J	Boot, Light Brown	Mouth, Light Brown
K	Boot, Dark Brown	Mouth, Gray
L	Mouth, Light Brown	Boot, Gray

Notes: Participants in Group II were shown the same 12 choice sets, but with the right and left packages reversed. In the case of choice set F, for example, participants in Group II saw the fetus warning on the left and the boot warning on the right. There were 52 participants in Group I and 45 participants in Group II.

C. Eye Tracking

Each participant was asked to sit at a distance of 65 cm in front of a 17-inch, 1280x1024-pixel LCD monitor of a Tobii T60 eye tracker (Tobii Technology 2011). While seated in front of the eye tracker, but before starting the task, participants underwent a standard calibration procedure (Tobii AB 2016). Immediately before the appearance of each of the 12 choice sets, the computer monitor showed only a fixation cross, centered on the screen, for 0.2 seconds. That way, participants were induced to fix their gaze at a predetermined point before looking at any details of the two packs in the choice set that was to appear next.

During the task, the eye tracker noninvasively recorded participants' eye movements at a sampling frequency of 60 Hz. The accompanying software (Tobii AB 2016) classified participants' eye movements into two types: *fixations* and *saccades*. A fixation corresponds to a state where the eye remains relatively still over a period of time, while a saccade corresponds to the rapid motion of the eye from one fixation to another. The classification is based on the velocity of the directional shifts of the eye (Salvucci and

Goldberg 2000). If the velocity is higher than 30 visual degrees per second, the eye movement is classified as a saccade. If the velocity is below this threshold, the eye movement was classified as part of a continuing fixation.

We further classified the participants' fixations according to their corresponding coordinates on the computer screen. In accordance with recommended practice (Holmqvist et al. 2011), we divided up the screen coordinates into five mutually exclusive *areas of interest*, specifically, the warning image, the warning text, the lateral text, the toxic-product symbol, and the brand name.



FIGURE 2. EYE FIXATIONS BY PARTICIPANT 30 ON CHOICE SET J

To illustrate the data acquisition process, Figure 2 above shows the eye fixations of participant 30 on choice set J, which appeared next to last in her 12-set task. Superimposed on the two-pack choice set is the standard representation for fixations and saccades (Salvucci and Goldberg 2000), where each fixation is a circle with diameter proportional to its duration, and where a connecting line represents a saccade. Her first fixation, lasting 180 milliseconds (ms), was on the image in the boot warning on the left. After a 50-ms saccade, her second fixation, lasting 580 ms, was on the image of the mouth warning at the right. Shown below the two packs is a time line of the participant's eye movements, where

each minor tick represents 100 ms. Her first fixation occurred 280 ms after choice set J appeared on the computer screen. Her second fixation occurred 510 ms into the task. At 1090 ms, the participant clicked her mouse, choosing the pack on the left as less risky, thus ending the task. In this choice set, which was introduced near the end of the experimental task, the participant fixated only once on each of two areas of interest on the computer screen and then made her choice. In other cases, participants fixated repeatedly on the same or different areas of interest, going back and forth between the two packs, before making a decision.

At the completion of the 12-set task, each participant was shown a diagram of his fixations for two of the choice sets and then asked to respond to three questions. Why do you think you looked at those areas, how did you select the pack that was less risky for your health, and do you think that this type of pack could make you smoke less? We refer to this final phase of our experiment as the “retrospective think-aloud.”

II. Theory

A. Additive and Lexicographic Utility Functions

Our objective in this section is not to break new theoretical ground, but to establish basic definitions that will guide the exposition below. In this study, our fundamental objects of choice are cigarette *packages*. Each package has two variable attributes, its *warning* and its *background color*. Let $W = \{Boot, Cadaver, Fetus, Mouth\}$ denote the finite set of warnings, where $w \in W$ is an arbitrary element of W . Let

$B = \{Gray, Light Brown, Dark Brown\}$ denote the finite set of background colors, where $b \in B$ is an arbitrary element of B . The set of packages is $X = W \times B$ with arbitrary element $x = (w, b) \in X$. Since W and B are finite, so is X .

In each successive computer screen, a cigarette smoker selects his preferred package from a binary *choice set* $S = \{x, x'\}$, where $x \neq x'$. We study whether these choices are consistent with a complete, transitive, binary preference relation on the set X of packages. Since the set X is finite, such a preference relation can be represented by a real-valued utility function $u: X \rightarrow \mathbb{R}$. That is, the smoker prefers the package x to the package x' if and only if $u(x) \geq u(x')$. Here, we use “prefers” as shorthand to mean that the smoker perceives x to be (weakly) less risky than x' . We do not imply that the smoker

would in fact purchase x from a given consumption bundle. In general, this utility representation admits the possibility that the smoker is indifferent between packages x and x' when $u(x) = u(x')$.

We specifically study whether the smoker's preferences can be represented by an *additive utility function* on the set X of packages. In that case, his utility function $u: X \rightarrow \bar{\mathbb{R}}$ can be decomposed as $u(x) = u(w, b) = u_W(w) + u_B(b)$, where $u_W: W \rightarrow \bar{\mathbb{R}}$ is a real-valued function on the set W of warnings and $u_B: B \rightarrow \bar{\mathbb{R}}$ is a real-valued function on the set B of background colors.

We also study whether the smoker's choices are consistent with a *lexicographic utility function* on the set X of packages. In that case, his utility function $u: X \rightarrow \bar{\mathbb{R}}$ satisfies the following conditions. There are separate real-valued functions $u_W: W \rightarrow \bar{\mathbb{R}}$ on the set W of warnings and $u_B: B \rightarrow \bar{\mathbb{R}}$ on the set B of background colors. Moreover, for any two packages $x = (w, b)$ and $x' = (w', b')$, $u(x) \geq u(x')$ when either:

(i) $u_W(w) > u_W(w')$; or (ii) $u_W(w) = u_W(w')$ and $u_B(b) \geq u_B(b')$.

The smoker with lexicographic utility thus compares two packages on the basis of their warnings and relies on their background colors only when the two packages have an equally risky warning. We don't formalize the opposite case where the smoker compares two packages based on their background colors and only on their warnings when they have an equally risky color. This restriction is based on our research findings, to be discussed below. As in the case of additive utility, the preferences implied by our lexicographic utility function are complete and transitive. We will refer to u_W as a *warning utility* and u_B as a *background-color utility*. Below, we use the terms "lexicographic utility" and "lexicographic preferences" interchangeably, recognizing that the former is in fact a representation of the latter on a finite set X of packages.

Not every additive utility function is lexicographic. But under certain conditions, lexicographic utility can be represented as an additive utility function. The following result is a special case of a more general result proved by Kohli and Jedidi (Kohli and Jedidi 2007). Assume that a smoker has lexicographic utility. A necessary and sufficient condition that his utility function can be expressed additively as $u(w, b) = u_W(w) + u_B(b)$ is: $\min_{w, w' \in W} \text{abs}(u_W(w) - u_W(w')) > \max_{b, b' \in B} \text{abs}(u_B(b) - u_B(b'))$. That is, the minimum absolute

difference between any two warning utilities exceeds the maximum absolute difference between any two background-color utilities.

To see how this condition works, suppose that $u_W(\text{Mouth}) = 0$, $u_W(\text{Fetus}) = 3$, $u_W(\text{Cadaver}) = 6$, and $u_W(\text{Boot}) = 9$, while $u_B(\text{Dark Brown}) = 0$, $u_B(\text{Light Brown}) = 1$, and $u_B(\text{Gray}) = 2$. Then $\min_{w, w' \in W} \text{abs}(u_W(w) - u_W(w')) = 3$, $\max_{b, b' \in B} \text{abs}(u_B(b) - u_B(b')) = 2$, and thus the condition holds. The additive utility function $u(w, b) = u_W(w) + u_B(b)$ in this example is in fact lexicographic.

Our characterization of the conditions under which lexicographic utility admits an additive representation suggests an empirical test for the existence of lexicographic preferences. First, we use the data on participants' binary choices to estimate the utilities u_W and u_B as parameters in a discrete choice model such as conditional logit with an additive utility specification. Second, we use the estimated values \hat{u}_W and \hat{u}_B to compute the statistic $Q = \min_{w, w' \in W} \text{abs}(\hat{u}_W(w) - \hat{u}_W(w')) - \max_{b, b' \in B} \text{abs}(\hat{u}_B(b) - \hat{u}_B(b'))$, and then employ bootstrap methods to test the one-sided null hypothesis that $Q \leq 0$.

Below, we also consider a variant of lexicographic utility in which the smoker is indifferent between two or more warnings or background colors. To see how this works, suppose that we change the very first warning utility in the foregoing example to $u_W(\text{Mouth}) = 3$, but leave all the remaining utilities unchanged. Then we could collapse the mouth and fetus warnings into a single attribute named *Mouth or Fetus* and write $u(\text{Mouth or Fetus}, b) = 3 + u_B(b)$ for any background color b . Kohli and Jedidi describe this variant as "satisficing lexicographic preferences."

B. Context-Dependent Preferences

There is a plethora of experimental evidence supporting the notion that individual preferences are context-dependent (Tversky and Simonson 1993). Most of the supporting data involves a demonstration that an individual's preference for x versus x' depends on the presence or absence of a third option x'' , sometimes referred to as a decoy (Trueblood et al. 2013, Rooderkerk, Van Heerde, and Bijmolt 2011). These experimental findings undermine the theoretical notion of "simple scalability" that an individual's binary

preference between x and x' depends solely on the comparative utilities $u(x)$ and $u(x')$ (Tversky 1972). Context-dependent effects provide one explanation why the behavior of participants in experimental settings does not always adhere to the weak axiom of revealed preference (WARP) (Hey 2001, Agranov and Ortoleva 2017). Apparently faced with the same choice set $S = \{x, x'\}$, the participant chooses x at one point in the experiment, but opts for x' later on. The resolution of the apparent contradiction is that some contextual element must have changed.

There are two natural contextual elements in our experimental design: the right-left orientation of the two packages in each binary choice set, and the sequential order of each choice set in the 12-set task. When the right-left positioning of a package on the computer screen influences the smoker's choices, we'll say that his preferences exhibit a *positioning effect* (Ryan, Krucien, and Hermens 2018). Such an effect may become important in a visual choice experiment when the subject has a systematic preference for objects in some parts of the visual field. When the order of presentation of a choice set influences the smoker's choices, we'll say that his preferences exhibit an *ordering effect*. Such an effect may become important when there is learning or fatigue during the course of the experiment (Campbell et al. 2015, Day et al. 2012, Czajkowsk, Giergiczny, and Greene 2014).

In our experimental setting, the *context* in which a package is displayed consists of its screen *position* and a sequential *order*. Let $M = \{Left, Right\}$ denote the set of screen positions with element $m \in M$, and let $N = \{1, \rightleftharpoons, 12\}$ denote the set of sequential orders with element $n \in N$. The extended set $Y = X \times M \times N$ with element $y = (x, m, n)$ thus contains all possible packages in all possible contexts. To take a concrete example, at the 11th computer screen in his experimental task, participant 30 was confronted with choice set J shown in Figure 2 above. This choice set contained the elements $\{y, y'\}$, where $y = (x, m, n)$ with $x = (Boot, Light Brown)$, $m = Left$, and $n = 11$, and where $y' = (x', m', n')$ with $x' = (Mouth, Light Brown)$, $m' = Right$, and $n' = 11$.

A complete, transitive binary preference relation on the finite set Y can likewise be represented as a real valued *context-dependent* mapping $v: Y \rightarrow \mathbb{R}$. In particular, we can consider an extension of the context-free additive utility model of the form

$v(y) = u(x) + u_M(m) + u_N(n)$, where $u(x) = u_W(w) + u_B(b)$. When $u_M(\text{Left}) \neq u_M(\text{Right})$, we have a positioning effect, and when $u_N(n) \neq u_N(n')$ for $n \neq n'$, we have an ordering effect. A context-free utility function has no positioning effect and no ordering effect.

Consider an extension of the additive utility model with positioning effects, but no ordering effects. A package x located on the left side of the screen has utility

$$u(x) + u_M(\text{Left}),$$

while a package x' located on the right has utility $u(x') + u_M(\text{Right})$.

For a smoker with a systemic preference for packages on the right, we have

$$u_M(\text{Right}) - u_M(\text{Left}) > 0.$$

In the case where $0 < |u(x) - u(x')| < u_M(\text{Right}) - u_M(\text{Left})$, the smoker would choose x over x' when x is on the left, but x' over x when x is on the right. Unless we account for positioning effects, we may conclude that the smoker's choices are unstable and violate WARP. Within the additive utility framework, we can similarly construct an example where failure to consider positioning effects would lead us to conclude that the smoker's choices violated transitivity. Failure to consider ordering effects can likewise result in apparent violations of WARP and transitivity.

III. Results

A. Stability of Preferences

We first address whether smokers' preferences for cigarette packages are stable without context effects and thus adhere to WARP. As before, we emphasize that the preferred cigarette pack is the one perceived as less risky. Figure 3 displays the choice sets E and K shown to the 52 participants assigned to Group I. The remaining 45 participants assigned to Group II were shown the same two choice sets, except that in our internal labeling scheme, the labels E and K were reversed.

Choice Set E



Choice Set K



FIGURE 3. CHOICE SETS E AND K SHOWN TO PARTICIPANTS IN GROUP I.

In Figure 3, choice set E shows the product $x = (\text{Mouth}, \text{Gray})$ on the left and the product $x' = (\text{Boot}, \text{Dark Brown})$ on the right. Choice set K shows the same products x and x' , except that the right-left orientation is reversed. Stable preferences without context effects would require that a participant consistently chose x or x' . As shown in row 1 of Table 3 below, 93 out of a total of 97 participants in our experiment gave stable responses, while only 4 (or 4.4 percent) showed apparent violations of WARP. Among the 93 participants with stable preferences, 89 (95.7%) consistently chose $(\text{Boot}, \text{Dark Brown})$ over $(\text{Mouth}, \text{Gray})$.

TABLE 3. TESTS OF STABILITY, ADDITIVE UTILITY, AND LEXICOGRAPHIC PREFERENCES

Test	Choice Sets	Package on Left	Package on Right	Condition§	N*
1	E	Mouth, Gray	Boot, Dark Brown	WARP	93
	K	Boot, Dark Brown	Mouth, Gray		
2	A	Foot, Dark Brown	Mouth, Dark Brown	Transitivity, Additive utility	97
	C	Boot, Light Brown	Foot, Light Brown		
	J	Mouth, Light Brown	Boot, Light Brown		
3	A	Mouth, Dark Brown	Foot, Dark Brown	Lexicographic preferences	89
	I	Foot, Gray	Mouth, Dark Brown		
4	E	Mouth, Gray	Boot, Dark Brown	Lexicographic preferences	88
	J	Boot, Light Brown	Mouth, Light Brown		
	K	Boot, Dark Brown	Mouth, Gray		
	L	Mouth, Dark Brown	Boot, Gray		
5	C	Boot, Light Brown	Foot, Light Brown	Transitivity, lexicographic preferences	91
	D	Fetus, Light Brown	Foot, Gray		
	F	Boot, Gray	Fetus, Gray		
6	B	Fetus, Dark Brown	Fetus, Light Brown	Lexicographic preferences	77
	G	Foot, Light Brown	Foot, Dark Brown		
1–6					63

§ WARP = Weak axiom of revealed preference. All conditions include context-free utility.

*N = Number of participants whose choices satisfied each specific test. There were 97 total participants. A total of 63 participants satisfied all 6 tests.

B. Transitivity of Preferences

We next address whether smokers' preferences for cigarette packs adhere to the axiom of transitivity in a context-free additive utility model. Figure 4 below displays the choice sets A, C and J shown to participants assigned to Group I. Again, the remaining participants in Group II were shown the same choice sets, except that the right-left positioning of the packages in each screen was reversed.

In choice set A, the smoker has to choose between (*Mouth, Dark Brown*) on the left and (*Cadaver, Dark Brown*) on the right. Under the context-free additive utility model $u(w, b) = u_w(w) + u_b(b)$, his choice will depend only on the relative values of $u_w(\textit{Mouth})$ and $u_w(\textit{Cadaver})$, as both packages have the same background color. In choice set C, the smoker has to choose between (*Cadaver, Light Brown*) on the left and (*Boot, Light Brown*) on the right. Again, under an additive utility model, his choice will

depend only on the relative values of $u_w(Cadaver)$ and $u_w(Boot)$, as both packages likewise have the same background color. Similarly, under an additive utility model, the smoker’s choice in set J will depend on the relative values of $u_w(Boot)$ and $u_w(Mouth)$.

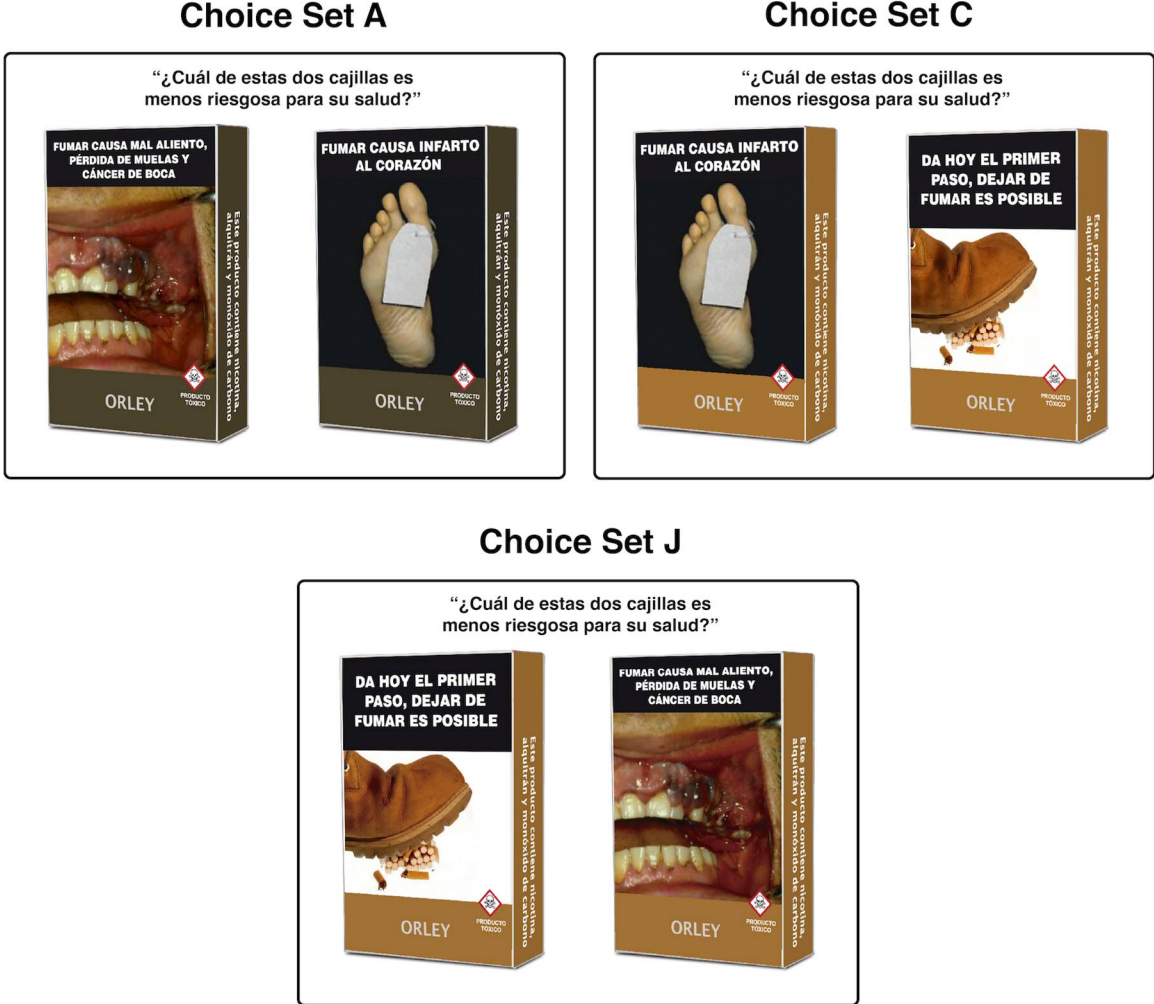


FIGURE 4. CHOICE SETS A, C AND J SHOWN TO PARTICIPANTS IN GROUP I.

Comparison of the smoker’s choices in sets A, C and J thus constitutes a test of transitivity. For example, if the smoker chooses the package on the right in set A, then $u_w(Cadaver) > u_w(Mouth)$, and if he chooses the package on the right in set C, then $u_w(Boot) > u_w(Cadaver)$. These two choices imply $u_w(Boot) > u_w(Mouth)$. Thus, transitivity would require that the smoker choose the package on the left in set J.

As shown in row 2 in Table 3 above, all 97 participants made choices among screens A, C and J that were consistent with transitivity in a context-free additive utility

model. The most common utility ordering was $u_w(\textit{Boot}) > u_w(\textit{Cadaver}) > u_w(\textit{Mouth})$, observed in 55 (56.7 percent) of participants. The second most common implied ordering was $u_w(\textit{Cadaver}) > u_w(\textit{Boot}) > u_w(\textit{Mouth})$, observed in 26 (26.8 percent) of participants, while the third most common was $u_w(\textit{Boot}) > u_w(\textit{Mouth}) > u_w(\textit{Cadaver})$, observed in 13 (13.4 percent).

We have thus verified that the smokers who participated in our experiment made choices among cigarette packages that adhered to the weak axiom of revealed preference and the transitivity property in a context-free additive utility model. Yet in the retrospective think-aloud task at the end of the experiment, many participants described the images of the fetus, cadaver and mouth tumor as frightening (“espantosa”), disgusting (“asquerosa”), and horrible (“horrible”) – in short, as aversive stimuli (Fehr and Rangel 2011). Said one participant, “The image of that mouth is so disgusting that I didn’t want to see it.” (“La imagen de la boca me da un asco que no quise verla.”) Said another, “The image of the baby impresses me. I can’t even look at it.” (“Me da mucha impresión la imagen del bebé. No la puedo ni mirar.”) Yet another said, “I think the image of the dead feet is strong, but the mouth is disgusting.” (“Yo creo que la imagen de los pies muertos es fuerte, pero la de la boca es asquerosa.”)

C. Lexicographic Preferences

We next test whether the choices of our participating smokers, which appeared to be consistent with an additive utility model, also conformed to lexicographic preferences. Our test was motivated in part by the statements of many participants in the retrospective think-aloud task that they took account of the background color only when both packages had the same warning. Said one participant, “When I saw two images were the same, I went with the lighter color.” (“Cuando veía dos imágenes iguales, me guiaba por el color más clarito.”) Said another, “I was guided by all the images, by the photo. In case they were the same, I focused on the colors and chose the lighter color.” (“Me guié en todas las imágenes, por la foto. En las que era igual, me fijaba en los colores y elegía el color más claro.”)

Figure 5 below shows the choice sets A and I displayed to the 52 participants in Group I. Choice set A compares (*Mouth, Dark Brown*) on the left with

(*Cadaver, Dark Brown*) on the right, while choice set I compares (*Cadaver, Gray*) on the left with (*Mouth, Dark Brown*) on the right. If a smoker with context-free lexicographic preferences chooses (*Cadaver, Dark Brown*) on the right in choice set A, he will choose (*Cadaver, Gray*) on the left in choice set I, even though the two packages have different background colors. Similarly, if he chooses (*Mouth, Dark Brown*) on the left in set A, he will also choose that package when it appears on the right in set I.

Any reversal would not constitute a violation of additive utility, but it would violate lexicographic preferences. For example, if a smoker with an additive utility function $u(w, b) = u_w(w) + u_b(b)$ chose (*Cadaver, Dark Brown*) on the right side in choice set A and (*Mouth, Dark Brown*) on the right side in choice set I, we can conclude only that $u_b(\text{Dark Brown}) - u_b(\text{Gray}) > u_w(\text{Cadaver}) - u_w(\text{Mouth}) > 0$. That is, the smoker trades off the warning and the background color in his perception of riskiness.



FIGURE 5. CHOICE SETS A AND I SHOWN TO PARTICIPANTS IN GROUP I.

As shown in row 3 in Table 3 above, 89 (91.8 percent) of our 97 participants made choices consistent with lexicographic utility. Among these 89 participants, 77 (86.5%) consistently chose the package with the cadaver warning in both sets.

Figure 6 shows the choice sets underlying an additional test of context-free lexicographic utility, based upon a comparison of the responses to choice sets E, J, K, and L, all of which paired a package with a boot warning to a package with a mouth warning. If

a smoker has lexicographic preferences, he will consistently choose either the package with the boot warning or the package with the mouth warning in all four choice sets.

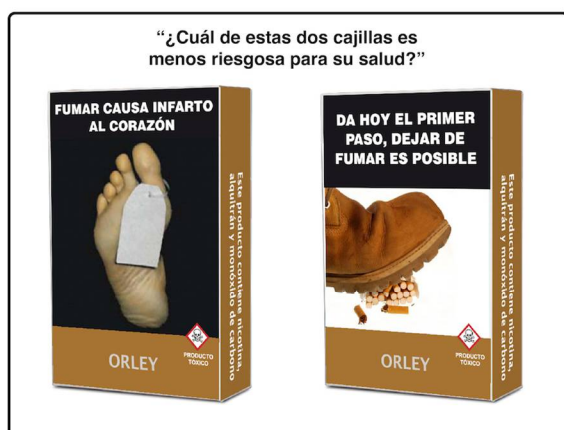


FIGURE 6. CHOICE SETS E, J, K AND L SHOWN TO PARTICIPANTS IN GROUP I.

As shown in row 4 in Table 3, 88 (90.7%) of the 97 smokers made choices among these four sets that were consistent with context-free lexicographic preferences. Among these 88 smokers, 85 (96.6%) preferred the packages with the boot warnings.

Figure 7 below shows the choice sets underlying yet another test of lexicographic preferences, based upon a comparison of the responses to choice sets C, D and F. If a smoker has lexicographic preferences, his choices among these three sets should display a transitive ordering among the boot, cadaver and fetus warnings that is independent of the background colors of the packages.

Choice Set C



Choice Set D



Choice Set F



FIGURE 7. CHOICE SETS C, D, AND F SHOWN TO PARTICIPANTS IN GROUP I.

As shown in row 5 in Table 3 above, 91 (or 93.8%) of the 97 smokers made choices among these three sets that were consistent with lexicographic preferences. The most common warning utility ordering was $u_w(\text{Boot}) > u_w(\text{Cadaver}) > u_w(\text{Fetus})$, observed in 54 (59.3 percent) of the 91 smokers who made choices consistent with lexicographic preferences among the three choice sets. The second most common warning ordering was $u_w(\text{Cadaver}) > u_w(\text{Boot}) > u_w(\text{Fetus})$, observed in 14 (15.4 percent), while the third most common was $u_w(\text{Boot}) > u_w(\text{Fetus}) > u_w(\text{Cadaver})$, observed in 13 (14.3 percent).

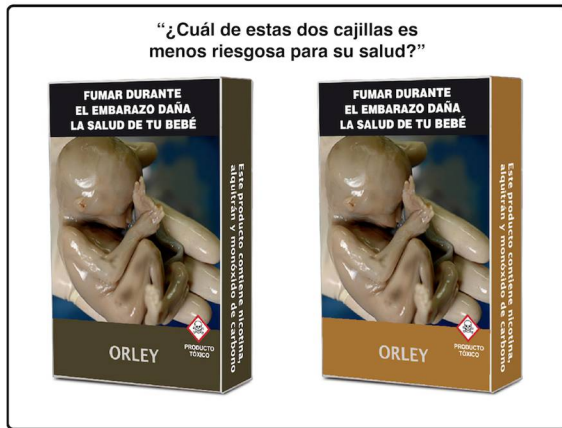
For each of the 91 smokers who displayed transitive lexicographic preferences in test 5, based on choice sets C, D and F, we can deduce an ordering among the warning utilities $u_w(\text{Boot})$, $u_w(\text{Cadaver})$, and $u_w(\text{Fetus})$. In particular, 68 (74.7%) of these 91

smokers made choices implying $u_w(\textit{Boot}) > u_w(\textit{Cadaver})$, while the remaining 23 made choices implying a reverse ordering of warning utilities. Moreover, each of these 91 smokers displayed transitive preferences in test 2 above, based on choice sets A, C and J. From those choices we can similarly deduce an ordering among the warning utilities $u_w(\textit{Boot})$, $u_w(\textit{Cadaver})$, and $u_w(\textit{Mouth})$. We can thus check whether the relative ordering of warning utilities $u_w(\textit{Boot})$ and $u_w(\textit{Cadaver})$ derived from test 5 is consistent with the relative ordering of the same utilities derived from test 2. In fact, the relative ordering was consistent in every one of the 91 smokers who displayed transitive lexicographic preferences in test 5.

Thus far, our tests of lexicographic preferences have focused on smokers' preferences over warnings as the dominant choice criterion. Test 6 addresses the preference ordering over background colors when both packages in a choice set have the same warning. Figure 8 draws a comparison between choice sets B and G. In set B, the smoker has to choose between $(\textit{Fetus}, \textit{Dark Brown})$ and $(\textit{Fetus}, \textit{Light Brown})$. In set G, the smoker instead has to choose between $(\textit{Cadaver}, \textit{Dark Brown})$ and $(\textit{Cadaver}, \textit{Light Brown})$. If a smoker with a context-free additive or lexicographic utility function chose $(\textit{Fetus}, \textit{Light Brown})$ on the right side in choice set B, he would be expected to choose $(\textit{Cadaver}, \textit{Light Brown})$ on the left side of choice set G.

As shown in the row 6 of Table 3, 77 (79.4%) of 97 participants gave responses consistent with either additive utility or lexicographic preferences. Among these 77 smokers, 67 (87.0%) displayed a consistent preference for light brown, while the remaining 10 displayed a consistent preference for dark brown.

Choice Set B



Choice Set G

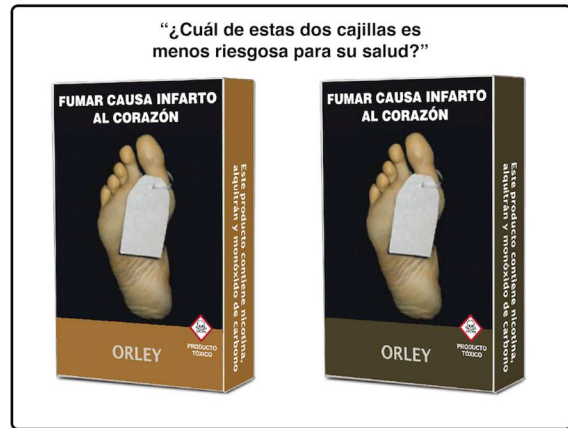


FIGURE 8. CHOICE SETS B AND G SHOWN TO PARTICIPANTS IN GROUP I.

The last row of Table 3 shows that 63 (64.9%) of the 97 smokers passed all six tests combined. Below, we use the term “violators” to characterize the 34 smokers who choices failed one or more of the six tests shown in the table. We use the term “non-violators” to characterize the 63 smokers whose choices passed all six tests.

D. Parametric Models

Table 4 below shows the results of conditional logit regressions on the $12 \times 97 = 1,164$ discrete choices made by our 97 participating smokers. Each of the three columns of parameter estimates in the table represents an additive utility model of the form $u_w(w) + u_b(b) + u_m(m)$, where $u_w(w)$ is the warning utility, $u_b(b)$ is the background-color utility, and $u_m(m)$ is the utility component that captures the participant’s preference for a package situated on one side of the computer screen. Below each parameter estimate is its estimated standard error. Also calculated for each model is the Q test statistic of the hypothesis that the estimated additive utility is consistent with lexicographic preferences, which we discussed above.

TABLE 4. CONDITIONAL LOGIT ADDITIVE UTILITY MODELS*

Independent variable	Model 1	Model 2	Model 3
Cadaver	-1.096 (0.177)	-1.112 (0.174)	-1.332 (0.175)
Fetus	-2.711 (0.266)		
Mouth	-2.864 (0.162)		
Fetus or Mouth		-2.835 (0.152)	-2.752 (0.149)
Light Brown	0.707 (0.178)	0.745 (0.164)	
Dark Brown	-0.507 (0.139)	-0.490 (0.134)	-0.825 (0.110)
Right Side	0.209 (0.082)	0.209 (0.082)	0.201 (0.081)
Q Statistic§	-1.061 (0.283)	-0.124 (0.224)	0.507 (0.141)

*All models had 1,164 observations on 97 participants. Numbers in parenthesis below each parameter estimate are standard errors.

§ $Q = \min_{w, w \in W} \text{abs}(\hat{u}_W(w) - \hat{u}_W(w)) - \max_{b, b \in B} \text{abs}(\hat{u}_B(b) - \hat{u}_B(b))$. We used bootstrap methods to compute the standard error of this statistic.

In model 1, the omitted reference category for the warnings was the boot, that is, $u_W(\text{Boot}) = 0$. Accordingly, the estimated utility component for the cadaver warning, relative to the boot warning, was $\hat{u}_W(\text{Cadaver}) = -1.096$ with standard error 0.177. The negative sign means that the cadaver warning was perceived as *more* risky than the reference boot warning. The omitted category for the background colors was gray, that is, $u_B(\text{Gray}) = 0$. Thus, the positive sign of $u_B(\text{Light Brown})$ means that the light brown color was perceived as *less* risky than the gray background color.

The utility component for left-sided positioning on the computer screen was set to $u_M(\text{Left}) = 0$. The estimate $\hat{u}_M(\text{Right}) = 0.209$ (standard error 0.082) indicates that for the sample as a whole, there was a significant right-sided positioning effect. The odds of choosing a package on the right side of the computer screen were an estimated 23.2 percent greater than the odds of choosing a package on the left ($\exp(\hat{u}_R) = 1.232$). The estimated utility parameters for the warnings and background colors were all different from zero at the significance level $p < 0.0003$. In a two-sided test, we could reject the null hypothesis

that $u_M(\textit{Right}) = 0$ at the level $p = 0.011$. Running the logit model with standard errors clustered by individual did not materially alter the significance levels.

In model 1, we could not reject the hypothesis that $u_W(\textit{Fetus}) = u_W(\textit{Mouth})$. (In a chi-squared test with one degree of freedom, $p = 0.574$.) The main reason for our inability to discriminate between these two warning utilities is the absence of a choice set in Table 2 that directly compares packages with the fetus and mouth warnings. To economize on parameters, we therefore ran model 2 under the restriction that smokers were indifferent between the two warnings, that is, $u_W(\textit{Fetus})$ and $u_W(\textit{Mouth})$ shared a common value $u_W(\textit{Fetus or Mouth})$. We retained the same reference categories as in model 1. This simplification did not significantly alter any of the estimated utility components.

Our finding of a significantly positive value of $\hat{u}_B(\textit{Light Brown})$ in both models 1 and 2 was unexpected, as it implied that the light brown background color was perceived as less risky than the gray background color. This finding went contrary to the original study among Australian smokers that included the three background colors studied here (Parr et al. 2011). In the retrospective think-aloud phase of the present experiment, many participants explicitly stated that they chose the lighter color (“el color más claro”) as less risky. Other experimental studies and cross-sectional surveys have supported the conclusion that lighter package colors are perceived as more healthful (Hammond et al. 2014, Wakefield, Germain, and Durkin 2008, Bansal-Travers et al. 2011, Doxey and Hammond 2011). In an earlier discrete choice experiment (Harris et al. 2017), we found that a package with a white background and a blue stripe was perceived as less risky.

As noted in our discussion of test 6 above, both choice sets B and G directly compared the light brown and dark brown background colors in packages with the same warning. There was no choice set directly comparing the gray background with either of these two colors. The data from choice sets B and G, we found, had significant leverage in the estimation of the parameters in model 1. When we included an interaction term for these two choice sets in model 1, the estimate for $u_B(\textit{Light Brown})$ was no longer statistically significant ($p = 0.344$). The dark brown color appears to have served as a decoy to enhance the attractiveness of the light brown alternative (Trueblood et al. 2013).

Accordingly, we tested an even more concise model 3, in which the light brown and gray background colors together served as the omitted category. In that model,

$u_B(\textit{Dark Brown})$ represented the utility component of dark brown relative to the remaining colors. In fact, $\hat{u}_B(\textit{Dark Brown})$ derived from model 3 was significantly different from the corresponding estimate derived from model 1 ($p = 0.0039$).

Table 4 also shows our calculations of our Q test statistic. The negative values of Q derived from models 1 and 2 were inconsistent with lexicographic preferences. However, the positive value of $Q = 0.507$ derived from model 3 was consistent with lexicographic utility for the entire sample of 97 smokers. The 95% confidence interval based on an asymptotically normal distribution, computed by bootstrap methods, was [0.229, 0.784]. A one-sided test of the null hypothesis that $Q \leq 0$ gave $p = 0.0002$. These results supported the conclusion that, at least for the more concise model 3, smokers had lexicographic preferences that could be represented by an additive utility function.

The models reported in Table 4 do not incorporate ordering effects. To test for such effects, we ran our three conditional logit additive utility models with additional right-hand-side variables representing the sequence order, either as a continuous variable or as fixed effects. We also tested interactions between sequence order and the other utility components of the model. In no case were the coefficients corresponding to the ordering effects significant.

As an additional test of ordering effects, we ran model 3 on subsets of the database partitioned by sequence number. We found no significant evidence of trend in the estimated coefficients during the course of the experiment. As an example, Figure 10 below shows the estimate of the warning utility $u(\textit{Cadaver})$ from model 3 in relation to the sequence ordering. We have collapsed the sequence number n into four groups (1–3, 4–6, 7–9, and 10–12) to reduce the dispersion of the data plot. Surrounding each point estimate is its 95 percent confidence interval. The figure shows no significant differences during the course of the experiment and thus provides evidence against ordering effects.

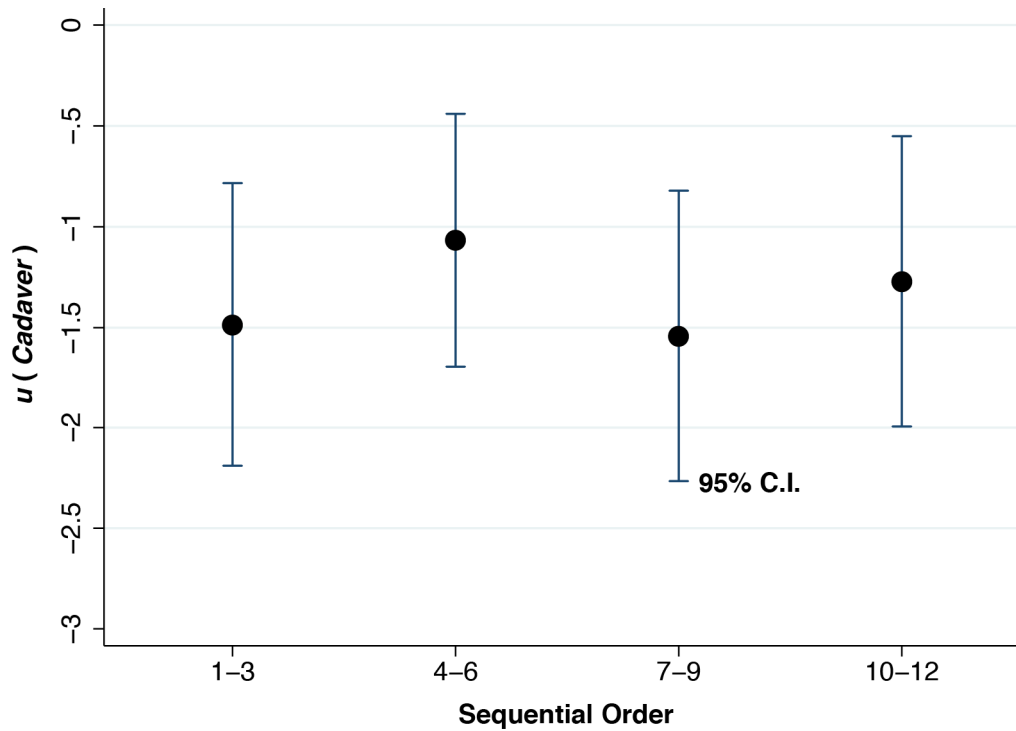


FIGURE 9. ESTIMATED UTILITY OF THE CADAVER WARNING RELATIVE TO THE BOOT WARNING IN RELATION TO SEQUENTIAL ORDER, BASED UPON MODEL 3

E. Response Time

As part of our tracking of the eye movements of participants, we monitored the time elapsed from the appearance of each choice set on the computer screen until the smoker clicked his mouse pointer on the cigarette pack he perceived as less risky. As we noted above, we imposed no time limit on the choice task. For all choice sets among all smokers, the mean *response time* was 4.408 seconds, with a median of 2.78 seconds and a range of 0.19 to 52.8 seconds.

Among the 12 choice sets described in Table 3, three sets (B, G and H) compared cigarette packages with identical warnings, while the remaining nine sets compared packages with distinct warnings. Figure 11 below shows the mean response time in relation to the screen sequence and the presence or absence of identical warnings.

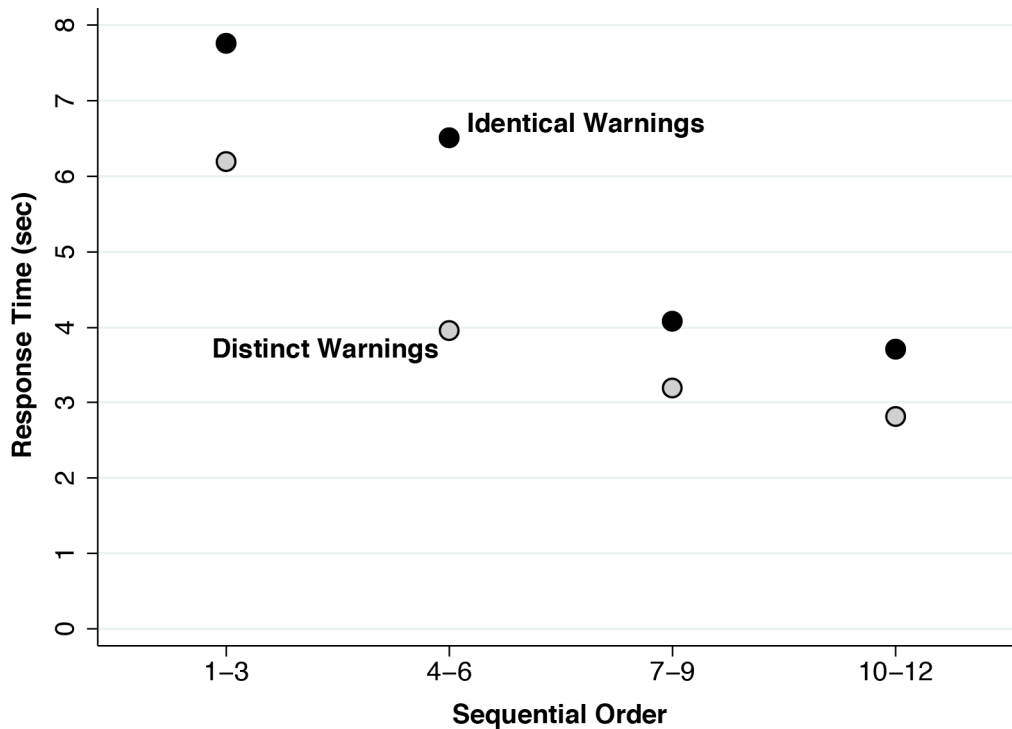


FIGURE 11. MEAN RESPONSE TIME IN RELATION TO SEQUENTIAL ORDER AND THE PRESENCE OR ABSENCE OF IDENTICAL PACKAGE WARNINGS IN THE CHOICE SET

As Figure 11 shows, there was an overall downward trend in the mean response time during the course of the experimental task. Whether early or late in the overall sequence of choice sets, participants spent more time reaching a decision when faced with a comparison of packages with identical warnings. By contrast, the mean response time had no relation to the presence or absence of identical background colors. In a linear regression of response time as a function of fixed effects for identical warnings, identical colors, screen sequence and each individual smoker, the estimated effect of identical warnings was 1.390 seconds (95% confidence interval [0.846, 1.934], $p < 0.0001$), while the corresponding parameter estimate for identical colors was -0.218 seconds (95% confidence interval $[-0.715, 0.280]$, $p = 0.391$). While Figure 11 suggests that the effect of identical warnings on response time diminished during the course of the experimental task, a fixed-effect regression did not show statistically significant interaction terms.

Our finding that identical package warnings – but not identical package colors – increased the response time is consistent with lexicographic preferences. If the smoker first evaluates the two package warnings and addresses differences in background color only

when the package warnings are identical, then we would expect a choice set with identical warnings to require more cognitive processing time.

For each of the 12 binary choice sets $\{x, x'\}$ in each of the two groups our experiment, we used the parameter estimates in model 3, shown in Table 4 above, to compute the quantity $\Delta\hat{u} = |\hat{u}(x) - \hat{u}(x')|$ as a measure of the divergence in utility between the two package alternatives. Since the calculation included the estimated positioning effect $\hat{u}_M(\text{Right})$, the estimated values of $\Delta\hat{u}$ for any given choice set differed between Groups I and II. When $\Delta\hat{u}$ is large, there is a strong preference for one of the alternatives, but when $\Delta\hat{u}$ is very small, the decision between the two packs is a close call. For each choice set and group, Figure 12 relates $\Delta\hat{u}$ to the mean response time.

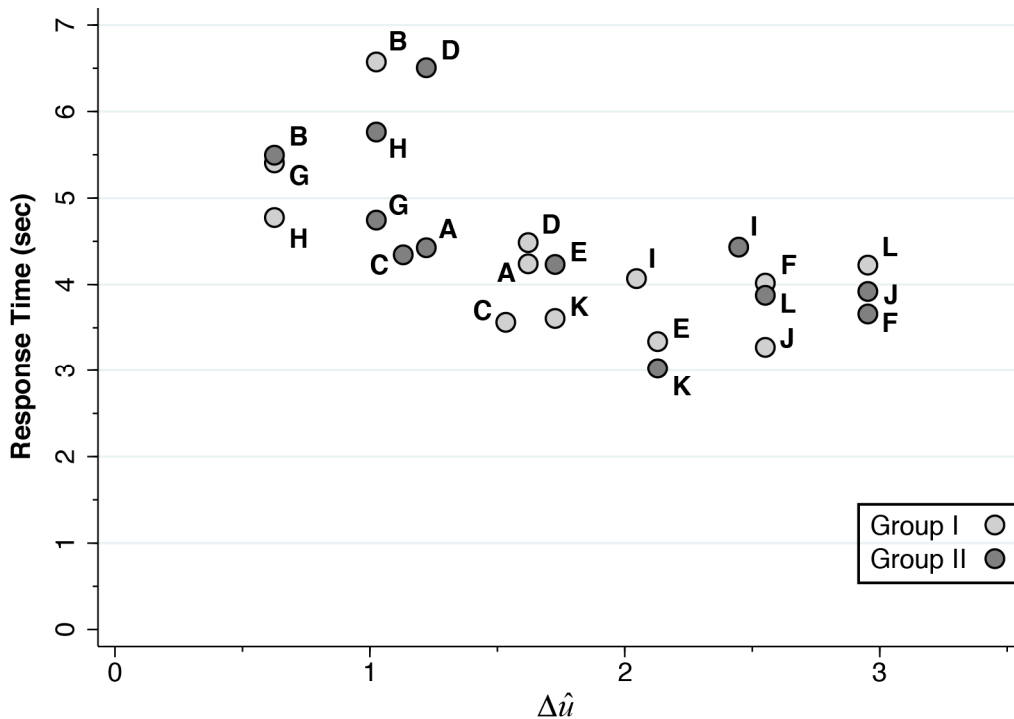


FIGURE 12. RELATION BETWEEN MEAN RESPONSE TIME AND ESTIMATED DIVERGENCE IN UTILITY AMONG 12 CHOICE SETS IN EACH OF THE TWO GROUPS

Figure 12 confirms that response time is inversely related to the estimated divergence in utility between the two package alternatives. The sets with the lowest values of $\Delta\hat{u}$ were B, G and H, precisely those in which the warnings were identical. The significant negative relation between the two variables was confirmed in a weighted least

squares regression, where the weights were the estimated inverse standard errors of the mean response time (estimated slope = -0.723 , $p = 0.001$). We also obtained a significant inverse relationship when we instead used models 1 and 2 to compute $\Delta\hat{U}$.

F. Eye Tracking

Figure 13 shows the initial fixation patterns of the 97 participants in relation to the presence or absence of identical warnings in the choice set. The figure shows two decision trees. The tree at the left corresponds to $3 \times 97 = 291$ choice sets with identical warnings, while the tree at the right corresponds to $9 \times 97 = 873$ choice sets with distinct warnings. These numbers appear inside the respective decision nodes at the root of each tree.

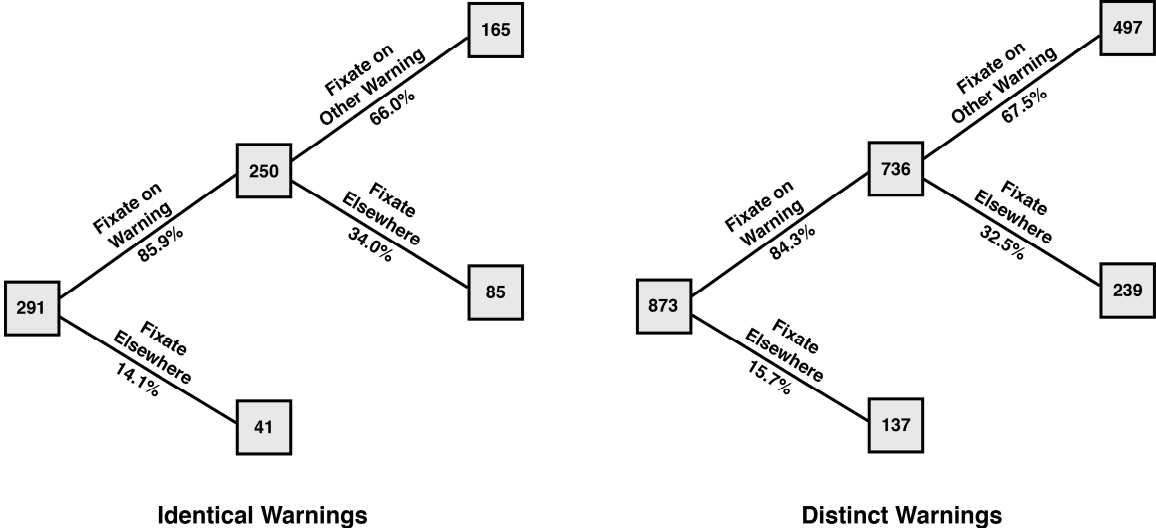


FIGURE 13. INITIAL FIXATION PATTERNS IN RELATION TO THE PRESENCE OR ABSENCE OF IDENTICAL PACKAGE WARNINGS IN THE CHOICE SET

Along the branches emanating from the decision node at the root of each tree, we show the proportions of choice sets in which participants did or did not initially fixate on a package warning, including the image or text. Thus, proceeding along the upper branch labeled “Fixate on Warning” in the tree on the left, we see that participants fixated initially on a warning in 250 (85.9%) of the 291 choice sets with identical warnings. In the tree on the right, proceeding along the corresponding branch labeled “Fixate on Warning,” participants fixated initially on a warning in 736 (84.3%) of 873 choice sets with distinct warnings. There was no significant difference between the two types of choice sets –

identical versus distinct warnings – in the probability of initially fixating on a package warning. (In a linear regression of initial fixation on a package warning as a function of the type of choice set and the sequence number n of the choice set, with fixed effects for each individual, the coefficient of the identical-warning type was 1.72% with $p = 0.417$.)

Proceeding along the branch labeled “Fixate on Warning” emanating from the root of each tree, we arrive at another decision node. The number of choice sets in which participants initially fixated on a warning (250 and 736, respectively) is shown inside each node. From that node, we determine the proportion of choice sets in which participants then fixated on the other package warning. While they may have fixated back and forth on the text and image of the first warning, the other warning had to be the next distinct area of interest to receive a fixation. If a smoker who initially fixated on areas within the first warning then fixated on (say) the brand name of either package, his choice would count instead toward the branch labeled “Fixate Elsewhere.” In the tree at the left, we find that participants next fixated on the other package warning in 165 (66.0%) of 250 choice sets. In the tree at the right, the proportion was 497 (67.5%) of 736 choice sets. There was no significant difference between the two types of choice sets in the conditional probability of fixating on the other package warning, given that a participant had initially fixated on one package warning ($p = 0.933$).

Figure 14 measures the mean number of fixations outside the two package warnings in relation to the sequence number of the choice set in the experimental task. Again, we distinguish between choice sets with and without identical package warnings.

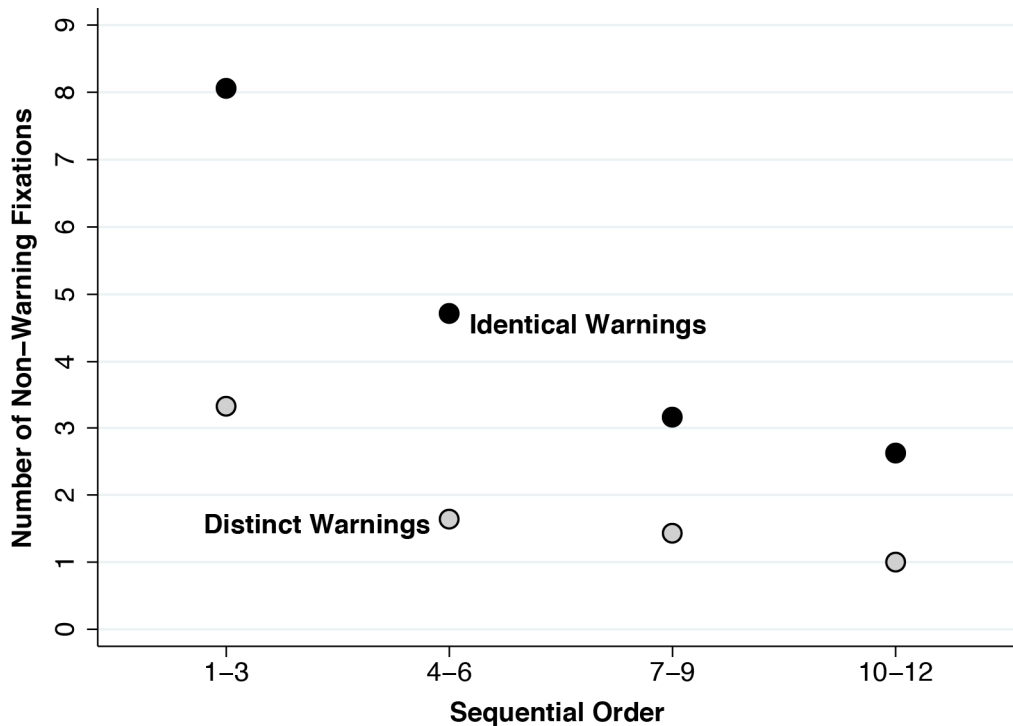


FIGURE 14. MEAN NUMBER OF FIXATIONS OUTSIDE OF PACKAGE WARNINGS IN RELATION TO SEQUENTIAL ORDER AND THE PRESENCE OR ABSENCE OF IDENTICAL PACKAGE WARNINGS IN THE CHOICE SET

As Figure 14 shows, participants made substantially more fixations on areas outside the package warnings when they were confronted with choice sets with identical warnings. The excess number of fixations declined during the course of the experimental task, as the sequence number n increased. In a linear regression of the number of non-warning fixations as a function of the type of choice set and sequence number, with fixed effects for each participant, all of the main effects and interactions were significant at the level $p < 0.05$.

We thus conclude that smokers' initial fixations were primarily focused on the warnings of the two packages in the binary choice set. These patterns of initial fixation were independent of the presence or absence of identical package warnings in the choice set. However, smokers confronted with choice sets containing identical warnings made significantly more subsequent fixations on package elements outside of the initial fixations on package warnings. These fixation patterns point to a cognitive process in which our smokers first assessed the two package warnings and, if they were identical, spent

additional effort distinguishing between the two packages on the basis of other package characteristics.

D. *Violators versus Non-Violators*

As summarized in Table 3, we performed six non-parametric tests of various conditions imposed by models of context-free additive or lexicographic utility. Of the 97 participants in our experiment, 63 (64.9%) passed all six tests, while 34 (35.1%) violated at least one of the six conditions. We termed the latter group the violators and the former group the non-violators. Here, we report on differences between these two groups.

We found no significant differences between violators and non-violators in mean age, the proportion of females, the proportion of students, or the proportion that attempted to quit in the past year. We found that 5 (14.7%) of 34 violators and 3 (4.8%) of 63 non-violators had no college education, but the difference was of borderline statistical significance (Pearson chi squared, $p = 0.089$). We calculated a two-component version of the Fagerström test for nicotine dependence (Heatherton et al. 1991), summing the respective 3-point scores for time to first cigarette and number smoked daily. The respective mean values were 1.65 for violators and 1.95 for non-violators, where the maximum score of 6 means the highest degree of addiction. The observed difference in means was not significant ($p = 0.307$).

We addressed the possibility that the apparent violations of additive utility or lexicographic preferences observed in a minority of our study participants were in fact a consequence of context-dependent preferences. To that end, we compared the total number of right-sided choices N_{Right} (where $12 \geq N_{Right} \geq 0$) among violators and non-violators of our six tests. The mean value of N_{Right} was 6.882 among the 34 violators, which was significantly greater than the mean value of 6.000 among the 63 non-violators. (In a two-sided t-test of group means with unequal variances, $p = 0.004$.) These findings suggested that the apparent violations of additive utility or lexicographic preferences were at least in part due to positioning effects.

For both violators and non-violators of our six tests, Figure 15 graphs the mean number of right-sided choices N_{Right} as a function of the sequence number of the choice set. Throughout the experimental task, from the first three to the last three choice sets, violators clicked more frequently on the right side of the computer screen than non-

violators. Moreover, we observed no significant relation between the sequence number n and the number of right-sided choices. (In a linear regression of study participants' binary right-versus-left choices against sequence number with fixed effects for each participant, $p = 0.385$.)

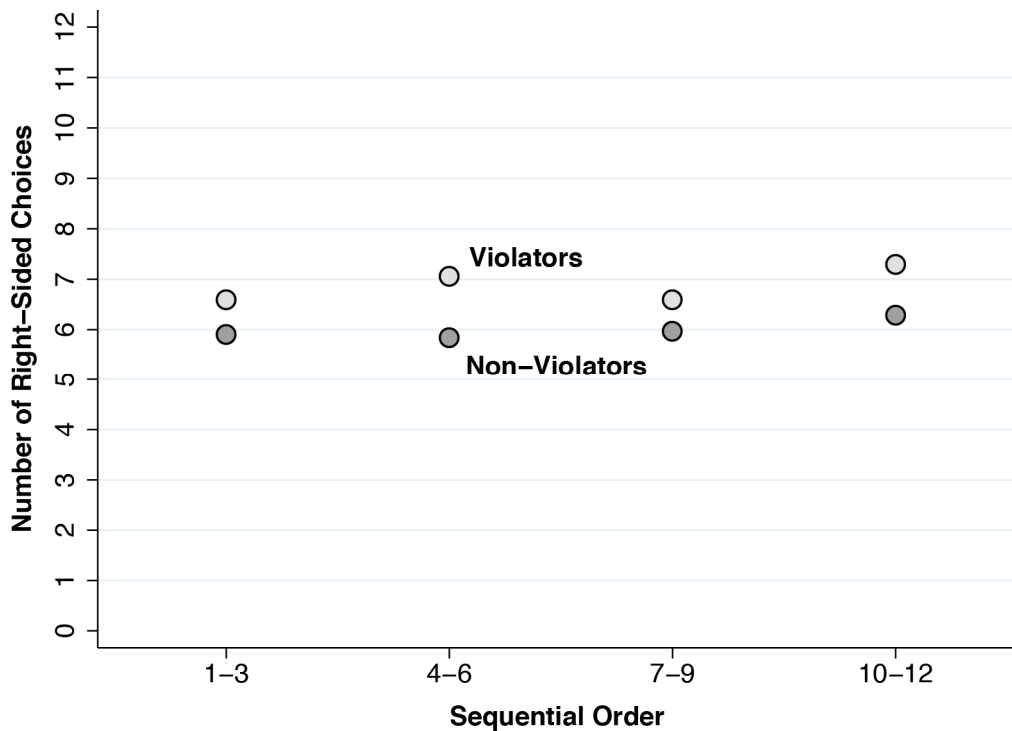


FIGURE 15. MEAN NUMBER OF RIGHT-SIDED CHOICES IN RELATION TO SEQUENTIAL ORDER FOR VIOLATORS AND NON-VIOLATORS OF TESTS 1–6

To determine whether ordering effects contributed to the apparent violations of additive and lexicographic utility, for each of the 97 participants we measured the absolute distance between the first and last choice set in each of our six tests. The underlying logic was that the larger the distance, the more susceptible would participants be to learning or fatigue effects. For example, for test 1 in Table 3, we calculated the absolute number of choice sets shown to each participant between set E and set K. The mean distances were 6.02 sets for violators and 6.75 sets for non-violators. (In a two-sided t-test of group means with unequal variances, $p = 0.692$.) Except for test 3, where the mean distance between sets A and I was 5.08 for violators and 6.75 for non-violators ($p = 0.093$), we found no significant differences between violators and non-violators.

These findings further support the conclusion that positioning effects, but not ordering effects, contributed to the apparent violations of at least one of the tests of additive utility and lexicographic preferences observed in 34 of the 97 smokers in our study.

Moreover, we found that violators had significantly longer response times than non-violators. Specifically, the mean response time was 5.470 seconds among the violators (standard error 0.329) and 3.848 seconds among the non-violators (standard error 0.137, $p < 10^{-5}$). We found that this significant difference in response times persisted even when we took our measure of divergence in utility $\Delta\hat{u}$ into account. In a weighted linear regression of response time as a function of $\Delta\hat{u}$ and an indicator variable for violator status, the effect of the latter variable was a 1.655-second increase in response time (standard error 0.333, $p < 10^{-6}$).

Finally, we found that violators made significantly more eye fixations than non-violators. Specifically, the mean number of fixations per choice set was 20.363 among the violators (standard error 1.257) and 12.447 among the non-violators (standard error 0.494, $p < 10^{-8}$). Moreover, violators spent more time on each fixation. Specifically, the mean duration per fixation was 178 ms among the violators (standard error 3.0 ms) and 169 ms among the non-violators (standard error 1.8 ms, $p = 0.018$). Linear regression models controlling for screen sequence gave nearly identical results.

III. Discussion

A. What We Found

In a discrete choice experiment, we asked 97 Uruguayan cigarette smokers to choose the less risky alternative among 12 pairs of cigarette packages with varying warnings and background colors. Some of the participants described the warnings as disgusting (“me da asco”) and frightening (“me da miedo”), and others had trouble even viewing them (“no la puedo ni mirar”). Nonetheless, nearly all were able to make choices that satisfied the basic axioms of rational decision-making, including transitivity and the weak axiom of revealed preference. What’s more, the majority of participating smokers made choices that satisfied six separate tests for the presence of lexicographic preferences.

These smokers, the tests showed, used the package warnings to decide which cigarette was less risky, and relied on background colors only to break ties.

To shed light on the cognitive processes underlying the participants' apparent lexicographic preferences, we focused on their differential responses to two types of binary choice sets – those in which both cigarette packs had identical warnings and those in which each pack had a distinct warning. We found that smokers spent significantly more time making a decision in the former case than in the latter. Utilizing the technique of eye tracking, we found that smokers initially approached both types of choice set in the same manner. They first fixed their gaze on the warnings on each of the two cigarette packs in the binary choice set shown on the computer monitor. Thereafter, their gaze patterns diverged. We observed significantly more fixations on non-warning elements of each cigarette package when the two packages had identical warnings. The observed sequence of eye fixations was consistent with a lexicographic choice strategy.

We investigated how contextual effects might explain why some smokers did not pass all six tests for lexicographic preferences. Even though the number of eye fixations and the total time to make a decision declined as the smoker was shown successive choice sets on the computer screen, there was no evidence of order-dependent preferences. We did find, however, that those smokers who did not pass all six tests for lexicographic preferences – whom we called the “violators” – exhibited a significantly greater preference for choosing packs on the right side of the computer screen.

B. What We Did Not Do

In this experiment, we asked each participant which cigarette package was less risky for his health. We did not ask him which product he would buy. We chose perception of the product's riskiness as an endpoint to avoid the complexities of interpretation inherent in intent to purchase, which would entail such additional intervening factors such as price and perceived cigarette taste.

When we subjected our discrete choice data to a series of increasingly concise conditional logit models, we found that the estimated parameters were consistent with a lexicographic utility function. This finding alone did not prove that the typical smoker in our sample in fact had lexicographic preferences. It merely supported the conclusion that the estimated weights in the typical smoker's additive utility function did not exclude lexicographic utility. Routine fitting of parametric utility models to discrete choice data

may mask the presence of underlying lexicographic decision strategies that can only be detected by non-parametric comparisons of the responses to individual choice sets, as we have done here.

When it comes to risk perception, we found evidence of lexicographic preferences over choice sets with two variable attributes: cigarette package warnings and background color. We did not establish that cigarette smokers have lexicographic preferences over all cigarette pack attributes.

Our finding that the violators of at least one of our six tests had a significantly greater preference for packages on the right side of the computer screen does not demonstrate that a right-sided positioning bias was the cause of the apparent violations of rationality. Their coincidence may be the result of some other unobserved factor that induces some smokers to make irrational choices and also to click on the right half of the screen.

While we found no evidence that smokers' preferences changed during the course of the experimental task, we did observe that the response time and the number of eye fixations declined with successive computer screens. One alternative explanation for the latter decline is learning. That is, as information on package attributes continues to accumulate in working memory, it becomes increasingly unnecessary to verify package attributes visually. An alternative explanation is fatigue (Campbell et al. 2015, Day et al. 2012, Czajkowski, Giergiczny, and Greene 2014).

We have focused on two specific sources of context-dependent preferences – what we have termed positioning effects and ordering effects. There may be others embedded in our smokers' choices, such as when the utility of a cigarette package depends on which package it is compared to. We observed suggestive evidence, in fact, that the utility component of the light brown background color $u_B(\textit{Light Brown})$ was enhanced when a package with a light brown background was juxtaposed against a package with a dark brown background. Such context-dependent preferences can be present even in experiments with only binary choice sets, but they may not be identifiable without imposing models with strong parametric assumptions.

We are hardly the first investigators to find that decision-makers use lexicographic strategies to choose among alternatives with multiple attributes (Colman and Stirk 1999, Tversky and Sattah 1979, Slovic 1975, Tversky, Sattah, and Slovic 1988, Rosenberger et al. 2003, Yee et al. 2007). Lexicographic heuristics are among the non-compensatory

strategies employed by decision makers to simplify the complexity of a decision problem (Ford et al. 1989, Gilbride and Allenby 2004, Montgomery and Svenson 1976, Tversky 1972, Gigerenzer and Goldstein 1996). While Ford and colleagues (Ford et al. 1989) suggested that non-compensatory strategies are more prevalent in complex tasks with large choice sets and multiple attributes, we found evidence of lexicographic preferences in a binary choice experiment with just two attributes. In our study, the set X of all cigarette packages, as well as the extended set Y of all cigarette packages in all possible contexts, are finite. Accordingly, the lexicographic strategy observed here is consistent with the maximization of a utility function representing a complete, transitive binary preference relation. In short, it adheres to the rules of rational decision-making.

Nor are we the first investigators to use the technique of eye tracking in the field of tobacco research. Others have used the technique to study which portions of the cigarette package individuals tend to focus on or avoid (Maynard et al. 2014, Krugman et al. 1994, Meernik et al. 2016, Kessels and Ruiters 2012, Shankleman et al. 2015, Munafo et al. 2011, Strasser et al. 2012). Our study follows a strand of research attempting to use the sequence of eye fixations – rather than simply the total number and duration of fixations – to elucidate information on search patterns (Russo and Rosen 1975). In that sense, our research strategy is part of a larger field of process tracing methods (Ford et al. 1989, Payne 1976).

C. How Can Addicted Smokers Be Rational?

We asked each participant how many cigarettes he smoked daily, and when did he smoke his first cigarette after arising in the morning. These two questions are components of the widely recognized Fagerström test for nicotine dependence (Heatherton et al. 1991), where they are each scored on a scale from 0 to 3. We summed these two scores to create an abbreviated Fagerström score ranging from 0 to 6. Based upon this abbreviated score, we found that those smokers who passed all six tests for rational, lexicographic preferences were no less addicted than those who failed at least one of the tests. Within the strict confines of their addiction, they acted rationally.

In a now famous 1978 article, Thomas Schelling described a model of choice in which “everyone behaves like two people, one who wants clean lungs and long life and another who adores tobacco... The two are in continual contest for control.” (Schelling 1978) The first of the two competing personas makes repeated attempts to quit smoking. In

our sample, 29 percent tried unsuccessfully to quit in the past year. The first persona also acknowledges that his behavior is dangerous. In our sample, 94 percent agreed that smoking causes lung cancer. The second persona, by contrast, engages in a compulsive, ritualized behavior often driven by strong cravings for a cigarette. In our sample, 60 percent had a cigarette within one hour of arising in the morning.

As part of the continual contest for control, the first persona employs a variety of strategies to dominate the second one. As Thaler and Shefrin noted in their seminal 1981 paper on the “two-self economic man,” “many smokers pay more for their cigarettes by buying them by the pack instead of the carton – it helps enforce a self-imposed ration such as one pack a day.” (Thaler and Schefrin 1981)

The second persona in turn employs an array of self-protective strategies to suppress the first one. In the retrospective think-aloud at the close of the experimental task, our smokers offered examples of such protective strategies. In an illustration of the strategy of self-exempting denial (Chapman, Wong, and Smith 1993), one participant commented, “The image is impressive, but I feel that these things won’t happen to me, so they don’t even affect me.” (“Te da una impresión esa imagen, pero creo que son cosas que a mi no me van a pasar, entonces ni me afectan.”) Another similarly commented, “What happens to me is that I’m not going to get these diseases, or at least for many years, so they don’t affect me.” (“Lo que me pasa es que esas enfermedades no me van a pasar, o por lo menos no dentro de muchos años, así que ni me afectan.”)

In an illustration of the masking strategy, one participant said, “I always cover the images, and that way I pass over them rapidly.” (“Yo siempre tapo las imágenes, por eso las pasé rápido.”) And another said, “When I buy a pack, I cover it. I don’t even look at it.” (Yo cuando compro una caja la tapo. Ni la miro.”) And still another said, “I always looked the other way.” (“Siempre miré para el otro lado.”)

Uruguay, a Spanish-speaking country in South America with about 3 million inhabitants, instituted a nationwide anti-smoking campaign in 2005 (Abascal et al. 2012, Triunfo, Harris, and Balsa 2016). In 2009, as part of the campaign, the government issued a “single presentation rule,” barring the marketing of multiple versions of the same brand, such as Silver, Blue, or Lights. The basis for such a policy was the presumption that such brand names misleadingly implied reduced harm.

From the viewpoint of the first persona in the Schelling-Thaler-Shefrin dual-self model, the smoker’s reliance on such colors as silver or blue is irrational. From the

viewpoint of the second persona, who operates under the constraints of addiction, color becomes a product attribute within a rational set of preference rules over cigarette packages. Here, we confirmed the findings of numerous other studies that smokers' perceptions of risk are in fact influenced by package colors (Hammond et al. 2014, Wakefield, Germain, and Durkin 2008, Bansal-Travers et al. 2011, Doxey and Hammond 2011, Harris et al. 2017).

From 2005–2009, during the course of Uruguay's tobacco control campaign, the proportion of pregnant women who were smoking at the start of their pregnancy dropped from about 25 to 15 percent. From 2007–2013, the proportion of women who had subsequently quit by the third trimester of pregnancy increased from about 15 to 43 percent. Yet among those women who quit during pregnancy and got pregnant again, nearly half had resumed smoking by the start of the subsequent pregnancy (Harris, Balsa, and Triunfo 2015). The oscillation between two states is consistent with Schelling's ongoing battle for control between the two personas. Unstable steady states were a predictable feature of the Becker-Murphy model (Becker and Murphy 1988).

D. Are Our Results Consistent with a Drift Diffusion Model of Choice?

Economists and neuroscientists have accumulated considerable evidence in support of a drift diffusion model of choice (Krajbich, Oud, and Fehr 2014, Ratcliff and McKoon 2008). In the tradition of satisficing models that offer alternatives to neoclassical utility maximization, the drift diffusion model (DDM) emphasizes the costs and benefits of acquiring additional information during the decision-making process. Most of the evidence supporting DDM comes from data on response times in binary choice experiments, although there are some results from eye tracking in binary and ternary choice settings (Krajbich and Rangel 2011, Krajbich, Armel, and Rangel 2010).

The basic idea underlying DDM is that the state of processing of information during decision-making is represented by a latent variable that drifts over time. In a binary choice setting, the decision-maker stops further information processing and makes a choice when the drifting state variable reaches one of two barriers. In particular, when a decision-maker is faced with a binary choice set $\{x, x'\}$, DDM predicts that his response time will be inversely related to $|u(x) - u(x')|$. If the characteristics of the two choices strongly

favor one of them, the velocity of drift toward the barrier is higher and the response time is shorter.

Our findings in Figure 12 support the DDM-predicted inverse relation between response time and the estimated divergence in utility $\Delta\hat{u} = |\hat{u}(x) - \hat{u}(x')|$, where $\hat{u}(x)$ and $\hat{u}(x')$ were derived from our conditional logit models of Table 4. Our results in Figure 11 go beyond this predicted inverse relation to specifically support a lexicographic strategy of information acquisition. We find that response times are longer when the two warnings are the same, so that $u_w(w) = u_w(w)$ and Δu thus narrows down to $|u_B(b) - u_B(b')|$. By contrast, we found that response times were no longer when the warnings were distinct and the colors were the same. The smokers in our experiment tended to stop searching for more information and click on the perceived less risky package once they encountered a distinction between the two package warnings.

In some experimental tests of the DDM, incorrect choices are also associated with increased response times (Ratcliff and McKoon 2008, Fudenberg, Strack, and Strzalecki 2017). In these experiments, there is generally an exogenous, observable standard to determine whether a choice is incorrect. Here, we have used the basic axioms of rational choice – WARP and transitivity – as well as the basic conditions underlying additive and lexicographic utility, as standards to test whether a choice is “incorrect.” We find that violators of one of these standards likewise have longer response times.

E. *What About Stochastic Preferences?*

In our theoretical section, we did not formally develop general stochastic versions of our additive or lexicographic utility functions. The fundamental question here is not whether one could do so, but whether such formalisms would add anything substantive to the data analysis and interpretation.

When we ran the conditional logit models reported in Table 4, we were implicitly assuming that the decision-maker’s utility function had an extra additive error term with an i.i.d. type I generalized extreme value distribution. We used the estimated logit coefficients to test the hypothesis that the resulting additive utility model was consistent with lexicographic preferences. We also used the logit specification to test the relation between response time and the estimated divergence in utility between the two choices.

The specification of a stochastic version of a more general – but not necessarily additive – lexicographic utility function would get us into unnecessary trouble. In our deterministic definition above, the package $\mathbf{x} = (w, b)$ is preferred to the package $\mathbf{x}' = (w', b')$, that is, $u(\mathbf{x}) \geq u(\mathbf{x}')$, when either: (i) $u_w(w) > u_w(w')$ or (ii) $u_w(w) = u_w(w')$ and $u_b(b) \geq u_b(b')$. Now suppose that we append an additive Fichnerian error term to the deterministic utility $u(\mathbf{x})$. If the distribution function of the error term has a density on the real line, then the condition $u_w(w) = u_w(w')$ in our deterministic definition would have a vanishing probability. Accordingly, we would instead have to specify a lexicographic semi-order (Tversky 1969, Manzini and Mariotti 2012) in which the decision maker is indifferent between two distinct package warnings if their corresponding utility difference is too small. In our notation, a smoker prefers $\mathbf{x} = (w, b)$ to $\mathbf{x}' = (w', b')$ when, for some number $\theta > 0$, either: (i) $u_w(w) - u_w(w') > \theta$; or (ii) $|u_w(w) - u_w(w')| \leq \theta$ and $u_b(b) \geq u_b(b')$. That leaves us with another parameter θ that, so far as we can determine, cannot be identified from our data.

One may object that, without a formalization of stochastic preferences, we cannot determine whether the violators identified in tests 1–6 in Table 3 actually made rational choices that were contaminated by random noise. Nor can we determine whether the non-violators actually made irrational choices that were masked by random noise. Such an objection, we think, borders on vacuous. The real issue is that the apparent violators did not make random errors. Instead, they systematically clicked on the right of the computer screen. One of our objectives in future research is to find out why.

F. *Where Do We Go From Here?*

While we reported on the number of fixations on the package warning, we did not distinguish between fixations on the image and text of the warning. In future research, we plan to study the relative importance of fixations on image versus text.

We found that the response time and the total number of fixations declined as the smoker was shown successive choice sets on the computer screen. In future research, we plan to study whether these observations are more consistent with learning or fatigue.

Studies utilizing eye tracking during a discrete choice experiment have not always found a strong relation between fixations and preferences (Balcombe et al. 2017). Researchers have recognized the difficulty of identifying the causal relations between fixation and choice solely from the data on the joint distribution of these two endogenous variables (Krucien, Ryan, and Hermens 2017, Shimojo et al. 2003). In top-down control of visual attention, preferences determine eye movements. In our context, the experimental participant tends to look at the cigarette package he eventually chooses. Under bottom-up control, by contrast, fixations drive preferences. Repeatedly looking at the warning enhances the probability that the pack will be chosen (Orquin and Mueller Loose 2013). In future research, we plan to develop instruments that distinguish between these two causal pathways in an attempt to further elucidate the eye-mind link (Reichle and Reingold 2013).

REFERENCES

- Abascal, W., E. Esteves, B. Goja, F. Gonzalez Mora, A. Lorenzo, A. Sica, P. Triunfo, and J. E. Harris. 2012. "Tobacco control campaign in Uruguay: a population-based trend analysis." *Lancet* 380 (9853):1575-82. doi: S0140-6736(12)60826-5 [pii] 10.1016/S0140-6736(12)60826-5.
- Agranov, Marina, and Pietro Ortoleva. 2017. "Stochastic Choice and Preferences for Randomization." *Journal of Political Economy* 125 (1):40-68.
- Aizaki, Hideo. 2012. "Basic functions for supporting an implementation of choice experiments in R." *Journal of Statistical Software* 50 (c02):1-24.
- Australian Government. 2011. "Tobacco Plain Packaging Regulations 2011: Select Legislative Instrument 2011 No. 263, pursuant to the Tobacco Plain Packaging Act 2011." *Federal Register of Legislative Instruments* F2011L02644 (<https://www.legislation.gov.au/Details/F2011L02644>).
- Balcombe, Kelvin, Iain Fraser, Louis Williams, and Eugene McSorley. 2017. "Examining the relationship between visual attention and stated preferences: A discrete choice experiment using eye-tracking." *Journal of Economic Behavior & Organization* 144 (December):238-257. doi: 10.1016/j.jebo.2017.09.023.
- Bansal-Travers, M., D. Hammond, P. Smith, and K. M. Cummings. 2011. "The impact of cigarette pack design, descriptors, and warning labels on risk perception in the U.S." *Am J Prev Med* 40 (6):674-82. doi: 10.1016/j.amepre.2011.01.021.
- Becker, G.S., and K.M. Murphy. 1988. "A theory of rational addiction." *Journal of Political Economy* 96 (4):675-700.
- Campbell, Danny, Marco Boeri, Edel Doherty, and W. George Hutchinson. 2015. "Learning, fatigue and preference formation in discrete choice experiments." *Journal of Economic Behavior & Organization* 119 (November):345-363.
- Chapman, S., W. L. Wong, and W. Smith. 1993. "Self-exempting beliefs about smoking and health: differences between smokers and ex-smokers." *Am J Public Health* 83 (2):215-9.
- CICT. 2016. *Tabaco : Advertencias con imágenes. Base de datos de MERCOSUR*. <http://www.cictmercosur.org/esp/galeria.php>: Comisión Intergubernamental para el Control del Tabaco del MERCOSUR.

- Colman, Andrew M., and Jonathan A. Stirk. 1999. "Singleton bias and lexicographic preferences among equally valued alternatives." *Journal of Economic Behavior & Organization* 40:337-351.
- Czajkowski, Mikolaj, Marek Giergiczny, and William H. Greene. 2014. "Learning and Fatigue Effects Revisited: Investigating the Effects of Accounting for Unobservable Preference and Scale Heterogeneity." *Land Economics* 9 (2):324-351.
- Day, Brett, Ian J. Bateman, Richard T. Carson, Diane Dupont, Jordan J. Louviere, Sanae Morimoto, Riccardo Scarpa, and Paul Wang. 2012. "Ordering effects and choice set awareness in repeat-response stated preference studies." *Journal of Environmental Economics and Management* 63 (1):73-91.
- Doxey, J., and D. Hammond. 2011. "Deadly in pink: the impact of cigarette packaging among young women." *Tob Control* 20 (5):353-60. doi: tc.2010.038315 [pii] 10.1136/tc.2010.038315.
- Fehr, Ernst, and Antonio Rangel. 2011. "Neuroeconomic Foundations of Economic Choice—Recent Advances." *Journal of Economic Perspectives* 25 (4):3-30. doi: doi=10.1257/jep.25.4.3.
- Ford, J. Kevin, Neal Schmitt, Susan L. Schechtman, Brian M. Hults, and M.L. Doherty. 1989. "Process tracing methods: contributions, problems and neglected research issues." *Organizational Behavior and Human Decision Processes* 43 (1):75-117.
- Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki. 2017. *Stochastic Choice and Optimal Sequential Sampling*. Working Paper, Available at Social Science Research Network: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2602927.
- Gigerenzer, Gerd, and Daniel G. Goldstein. 1996. "Reasoning the fast and frugal way: Models of bounded rationality." *Psychological Review* 103 (4):650-669. doi: 10.1037/0033-295X.103.4.650
- Gilbride, Timothy J., and Greg M. Allenby. 2004. "A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules." *Marketing Science* 23 (3):391-406. doi: 10.1287/mksc.1030.0032.
- Gruber, J., and B. Koszegi. 2001. "Is addiction rational? Theory and evidence." *Quarterly Journal of Economics* 116 (4):1261-1303.

- Hammond, D., C. White, W. Anderson, D. Arnott, and M. Dockrell. 2014. "The perceptions of UK youth of branded and standardized, 'plain' cigarette packaging." *Eur J Public Health* 24 (4):537-43. doi: 10.1093/eurpub/ckt142.
- Harris, J. E., G. Ares, M. Gerstenbluth, L. Machin, and P. Triunfo. 2017. "Impact of plain packaging of cigarettes on the risk perception of Uruguayan smokers: an experimental study." *Tobacco Control*. doi: 10.1136/tobaccocontrol-2017-053639.
- Harris, J. E., A. I. Balsa, and P. Triunfo. 2015. "Tobacco control campaign in Uruguay: Impact on smoking cessation during pregnancy and birth weight." *J Health Econ* 42:186-96. doi: 10.1016/j.jhealeco.2015.04.002.
- Heatherton, T.F., L.T. Kozlowski, R.C. Frecker, and K.O. Fagerstrom. 1991. "The Fagerstrom test for nicotine dependence: a revision of the Fagerstrom Tolerance Questionnaire." *British Journal of Addiction* 86 (9):1119-1127.
- Hey, John D. 2001. "Does Repetition Improve Consistency?" *Experimental Economics* 4 (1):5-54.
- Holmqvist, Kenneth, Marcus Nyström, Richard Anderson, Richard Dewhurst, Halszka Jarodzka, and Joost van de Weijer. 2011. *Eye Tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.
- Johnson, F. R., B. Kanninen, M. Bingham, and S. Ozdemir. 2007. "Experimental Design for Stated Choice Studies (Chapter 7)." In *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice*, edited by B.J. Kanninen, 159-202. Dordrecht: Springer.
- Kessels, L. T., and R. A. Ruiter. 2012. "Eye movement responses to health messages on cigarette packages." *BMC Public Health* 12:352. doi: 10.1186/1471-2458-12-352.
- Kohli, Rajeev, and Kamel Jedidi. 2007. "Representation and Inference of Lexicographic Preference Models and Their Variants." *Marketing Science* 26 (3):380-399.
- Krajbich, I., C. Armel, and A. Rangel. 2010. "Visual fixations and the computation and comparison of value in simple choice." *Nat Neurosci* 13 (10):1292-8. doi: 10.1038/nn.2635.
- Krajbich, I., and A. Rangel. 2011. "Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions." *Proc Natl Acad Sci U S A* 108 (33):13852-7. doi: 10.1073/pnas.1101328108.

- Krajbich, Ian, Bastiaan Oud, and Ernst Fehr. 2014. "Benefits of Neuroeconomic Modeling: New Policy Interventions and Predictors of Preference." *American Economic Review Papers and Proceedings* 104 (5):501-506.
- Krucien, Nicolas, Mandy Ryan, and Frouke Hermens. 2017. "Visual attention in multi-attributes choices: What can eye-tracking tell us?" *Journal of Economic Behavior & Organization* 135 (March):251–267.
- Krugman, Dean M., Richard J. Fox, James E. Fletcher, Paul M. Fischer, and Tina H. Rojas. 1994. "Do adolescents attend to warnings in cigarette advertising? An eye-tracking approach." *Journal of Advertising Research* (Nov/Dec):39-52.
- Lahey, Joanna N., and Douglas Oxley. 2016. "The Power of Eye Tracking in Economics Experiments." *American Economic Review: Papers & Proceedings* 106 (5):09–313.
- Manzini, Paola, and Marco Mariotti. 2012. "Choice by lexicographic semiorders." *Theoretical Economics* 7 (1):1-23. doi: 10.3982/TE679.
- Maynard, O. M., A. Attwood, L. O'Brien, S. Brooks, C. Hedge, U. Leonards, and M. R. Munafo. 2014. "Avoidance of cigarette pack health warnings among regular cigarette smokers." *Drug Alcohol Depend* 136:170-4. doi: 10.1016/j.drugalcdep.2014.01.001.
- Meernik, C., K. Jarman, S. T. Wright, E. G. Klein, A. O. Goldstein, and L. Ranney. 2016. "Eye Tracking Outcomes in Tobacco Control Regulation and Communication: A Systematic Review." *Tob Regul Sci* 2 (4):377-403. doi: 10.18001/TRS.2.4.9.
- Montgomery, Henry, and Ola Svenson. 1976. "On decision rules and information processing strategies for choices among multiattribute alternatives " *Scandinavian Journal of Psychology* 17 (1):283-291.
- Moodie, C., J. Hoek, J. Scheffels, K. Gallopel-Morvan, and K. Lindorff. 2018. "Plain packaging: legislative differences in Australia, France, the UK, New Zealand and Norway, and options for strengthening regulations." *Tob Control*. doi: 10.1136/tobaccocontrol-2018-054483.
- Munafo, M. R., N. Roberts, L. Bauld, and U. Leonards. 2011. "Plain packaging increases visual attention to health warnings on cigarette packs in non-smokers and weekly smokers but not daily smokers." *Addiction* 106 (8):1505-10. doi: 10.1111/j.1360-0443.2011.03430.x.

- Orquin, Jacob L., and Simone Mueller Loose. 2013. "Attention and choice: A review on eye movements in decision making." *Acta Psychologica* 144 (1):190-206.
- Parr, Victoria, Tan Benits, Patrick Eli, and Kyle Miller. 2011. *Market Research to Determine Effective Plain Packaging of Tobacco Products*. Sydney, Australia: GfK Blue Moon.
- Payne, John W. 1976. "Task complexity and contingent processing in decision making: An information search and protocol analysis." *Organizational Behavior and Human Performance* 16 (2):366-387.
- Ratcliff, Roger, and Gail McKoon. 2008. "The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks." *Neural Computation* 20 (4):873–922.
- Reichle, E. D., and E. M. Reingold. 2013. "Neurophysiological constraints on the eye-mind link." *Front Hum Neurosci* 7:361. doi: 10.3389/fnhum.2013.00361.
- Reutskaja, Elena, Rosemarie Nagel, Colin F. Camerer, and Antonio Rangel. 2011. "Search Dynamics in Consumer Choice under Time Pressure: An Eye- Tracking Study." *American Economic Review* 101 (2):900-926.
- Rooderkerk, Robert P., Harald J. Van Heerde, and Tammo H.A. Bijmolt. 2011. "incorporating Context effects into a Choice model." *Journal of Marketing Research* 48 (4):767-780.
- Rosenberger, Randall S., George L. Peterson, Andrea Clarke, and Thomas C. Brown. 2003. "Measuring dispositions for lexicographic preferences of environmental goods: integrating economics, psychology and ethics." *Ecological Economics* 44 (1):63-76.
- Russo, J.E., and L.D. Rosen. 1975. "An eye fixation analysis of multialternative choice." *Memory & Cognition* 3 (3):267-276.
- Ryan, M., N. Krucien, and F. Hermens. 2018. "The eyes have it: Using eye tracking to inform information processing strategies in multi-attributes choices." *Health Econ* 27 (4):709-721. doi: 10.1002/hec.3626.
- Salvucci, Dario D., and Joseph H. Goldberg. 2000. "Identifying fixations and saccades in eye-tracking protocols." In *ETRA '00 Proceedings of the 2000 symposium on Eye tracking research & applications* edited by Andrew T. Duchowski, 71-78. New York: ACM.
- Schelling, Thomas C. 1978. "Egonomics, or the art of self-management." *American Economic Review Papers and Proceedings* 68 (2):290-294.

- Shankleman, M., C. Sykes, K. L. Mandeville, S. Di Costa, and K. Yarrow. 2015. "Standardised (plain) cigarette packaging increases attention to both text-based and graphical health warnings: experimental evidence." *Public Health* 129 (1):37-42. doi: 10.1016/j.puhe.2014.10.019.
- Shimojo, S., C. Simion, E. Shimojo, and C. Scheier. 2003. "Gaze bias both reflects and influences preference." *Nat Neurosci* 6 (12):1317-22. doi: 10.1038/nn1150.
- Slovic, Paul. 1975. "Choice between equally valued alternatives " *Journal of Experimental Psychology: Human Perception and Performance* 1 (3):280-287. doi: 10.1037/0096-1523.1.3.280.
- Strasser, A. A., K. Z. Tang, D. Romer, C. Jepson, and J. N. Cappella. 2012. "Graphic warning labels in cigarette advertisements: recall and viewing patterns." *Am J Prev Med* 43 (1):41-7.
- Thaler, Richard H., and H.M. Schefrin. 1981. "An economic theory of self-control." *Journal of Political Economy* 89 (21):392-406.
- Tobii AB. 2016. *Tobii Studio User's Manual Version 3.4.5*: <https://www.tobiiipro.com/siteassets/tobii-pro/user-manuals/tobii-pro-studio-user-manual.pdf>.
- Tobii Technology. 2011. *Tobii T60 & T120 Eye Tracker: User Manual Revision 4.0*. Stockholm: Tobii Technology AB, [https://www.sllc.com/images/uploads/Tobii_T120_Eye_Tracker - User Manual.pdf](https://www.sllc.com/images/uploads/Tobii_T120_Eye_Tracker_-_User_Manual.pdf).
- Triunfo, P., J. Harris, and A. Balsa. 2016. "[Evaluation of Uruguay's antismoking campaign: progress and challenges at ten years]." *Rev Panam Salud Publica* 40 (4):256-262.
- Trueblood, J. S., S. D. Brown, A. Heathcote, and J. R. Busemeyer. 2013. "Not just for consumers: context effects are fundamental to decision making." *Psychological Science* 24 (6):901-8. doi: 10.1177/0956797612464241.
- Tversky, Amos. 1969. "Intransitivity of preferences." *Psychological Review* 76 (1):31-48. doi: 10.1037/h0026750
- Tversky, Amos. 1972. "Elimination by aspects: A theory of Choice." *Psychological Review* 79 (4):281-299. doi: DOI:10.1037/h0032955
- Tversky, Amos, and Shmuel Sattah. 1979. "Preference trees." *Psychological Review* 86:542-573.

- Tversky, Amos, Shmuel Sattah, and Paul Slovic. 1988. "Contingent weighting in judgment and choice." *Psychological Review* 95:371-384.
- Tversky, Amos, and Itamar Simonson. 1993. "Context-depedent preferences." *Management Science* 39 (10):1179-1189.
- U.S. Department of Health and Human Services. 1988. *The Health Consequences of Smoking - Nicotine Addiction: A Report of the Surgeon General*. Rockville MD: Centers for Disease Control, Office on Smoking and Health.
- Wakefield, M. A., D. Germain, and S. J. Durkin. 2008. "How does increasingly plainer cigarette packaging influence adult smokers' perceptions about brand image? An experimental study." *Tob Control* 17 (6):416-21. doi: 10.1136/tc.2008.026732.
- Yee, Michael, Ely Dahan, John R. Hauser, and James Orlin. 2007. "Greedoid-Based Noncompensatory Inference " *Marketing Science* 26 (4):532-549.