

研究論文

網路調查的注意力檢測對題項 回應時間與問卷完填的影響*

于若蓉** 杜素豪***

摘要

近二十年來，網路調查有益趨流行的趨勢。由於網路調查仰賴受訪者自填問卷，許多網路調查會在問卷中加入注意力檢測題，以偵測受訪者的專注程度。儘管注意力檢測題已廣被應用，嚴謹的方法研究仍不多見。特別是注意力檢測題對受訪者後續作答行為、退出問卷機率的影響，研究極少。在這項研究中，針對機率樣本進行網路調查實驗，設計了兩個注意力檢測題，並區分前、後兩個放置位置，再加上未設有注意力檢測題的對照組，總共區分為五組。藉由蒐集到的問卷資料、周邊資料，分析注意力檢測題對退出問卷機率、題項回應時間的影響效果。分析結果顯示，注意力檢測題對退出問卷機率並無顯著的影響效果。但在題項回應時間方面，放置於問卷後半部的注意力檢測題，會使得後續題項的回應時間顯著增加；放置於前半部的注意力

* 本研究是在科技部「網路調查的注意力檢測、資料品質與完訪結果」專題研究計畫（MOST 109-2410-H-001-016-SSS）經費補助下完成，特此申謝。此外，作者感謝三位審查人提供寶貴意見。然文責概由作者自負。

** 通訊作者。中央研究院人文社會科學研究中心調查研究專題中心研究員。臺北市南港區 11529 研究院路二段 128 號，電話：02-27898195，E-mail: yurr@gate.sinica.edu.tw。

*** 中央研究院人文社會科學研究中心調查研究專題中心研究員。E-mail: suhao@gate.sinica.edu.tw。

檢測題，則沒有顯著影響。這些結果顯示，如果研究者想在網路調查中藉由注意力檢測題提升受訪者的專注力，將注意力檢測題放在問卷後半部較前半部來得有效。

關鍵詞：注意力檢測、網路調查、調查實驗、題項回應時間、問卷完填

The Effects of Attention Checks on Item Response Latency and Questionnaire Completion

Ruoh-Rong Yu*, Su-Hao Tu**

ABSTRACT

Online surveys have become more prevalent due to the widespread use of the Internet and mobile devices. Because online survey data are collected by self-administered questionnaires, data quality is highly related to respondents' attentiveness. Attention checks have been developed by researchers to identify inattentive respondents in online surveys by means of asking the respondents to respond according to the instructions of the question rather than their true answer.

Even though passing an attention check was found to be associated with the respondent's overall data quality, several researchers have argued that a respondent may feel uncomfortable or not trusted when he or she encounters an attention check question. However, only a few studies, mostly based on non-probability samples, have examined whether attention checks affect respondents' subsequent behaviors. By conducting a survey experiment centered on a probability-based online panel, this study analyzes the effects of attention checks on the

* Corresponding Author. Research Fellow, Center for Survey Research, Research Center for Humanities and Social Sciences, Academia Sinica. No. 128, Academia Road, Section 2, Nankang, Taipei 11529, Taiwan. Tel: +886-2-27898195. E-mail: yurr@gate.sinica.edu.tw.

** Research Fellow, Center for Survey Research, Research Center for Humanities and Social Sciences, Academia Sinica. E-mail: suhao@gate.sinica.edu.tw.

likelihood of dropping out from an online survey and the response latency to subsequent items.

The sample used in this study is based on an online panel constructed by the Center for Survey Research, Academia Sinica, Taiwan. To examine whether the design and location of attention check questions matter for the respondents' subsequent behaviors, we adopted a two-by-two experimental design in which there are two designs and two locations (former half and latter half) for attention check questions. This sample was randomly assigned to these four treatment groups and one additional control group (with no attention check questions).

The theme of the online survey is the attitudes toward surveys and data. The online survey was conducted from April 19 to May 3, 2021. Invitation emails were sent successfully to 7,400 subjects. Among the 3,875 respondents, 3,770 completed the questionnaires. A logit model and Cox proportional hazard model were used to analyze the effects of attention checks on the probability of withdrawal. A multilevel mixed-effects generalized linear model was applied to estimate the impact of an attention check on response latency of subsequent items. Possible outliers of response latency were identified and treated in different ways as a robustness check.

The results indicate that adding an attention check to a questionnaire does not matter for the probability of withdrawal. This finding is intact with regards to the design or location of attention checks. Our findings also reveal that attention checks located in the latter half of a questionnaire have positively significant effects on the response latency of subsequent items, yet the attention checks placed in the former half are insignificant. The findings on response latency are robust to different ways of identifying and treating outliers. In sum, our results indicate that even though attention checks are not associated with the probability of completing a questionnaire, their locations are important for whether respondents' attentiveness can be effectively elevated.

Despite our finding that attention checks did not provoke negative

behaviors such as quitting the questionnaire or speeding up the response time, a preliminary analysis on response latency to attention checks reveals that the respondents who failed an attention check spent more time in responding to the attention check than those who passed. Responding behaviors to the attention check itself deserve further exploration in future research.

Keywords: attention checks, online survey, survey experiment, response latency, questionnaire completion

一、研究背景與目的

近二十年來，隨著網路、行動裝置的普及，網路調查有益趨流行的趨勢。相較於面訪、電訪等調查模式，網路調查具有低成本、即時性等優勢。由於網路調查是採用自填問卷的方式進行資料蒐集，可免於由訪員進行問卷訪談可能產生的訪員效應（interviewer effects）問題（Biemer and Stokes 1989）。相較由訪員進行問卷訪談的面訪、電訪模式，對於較敏感（sensitive）或較可能存在社會讚許（social desirability）顧慮的題項，網路調查的受訪者在答題時較可能根據真實的情況、想法作答，而有較低的回答誤差（Kreuter et al. 2008）。然而，由於網路調查沒有訪員提問、協助釋疑，受訪者只能仰賴自己閱讀、解讀題目，萃取答題所需的資訊，再由所萃取的資訊決定答案，並自己進行填答（Tourangeau et al. 2000）。相較由訪員進行訪談的問卷資料蒐集模式，網路調查對受訪者認知能力的要求較高，受訪者讀取題目、思考答案、填寫答案所需要花費的心力也較多。如果受訪者在填答網路自填問卷時不夠專注，可能無法正確解讀題意、跳過資訊萃取或思考答案的過

程，或在答題時誤選答案。此外，相較由訪員進行問卷訪談的調查模式，網路調查的受訪者在思考答案、提供答案時，可能會儘量花最少的力氣，採用最低滿意程度（*satisficing*）而非最適的（*optimal*）作答策略。¹ 採用最低滿意程度作答策略的受訪者，在面對問卷題目時，不會費心瞭解題目，在填答時也可能不會費神思索，想出最符合想法或實際情況的答案（Krosnick 1991）。更極端的情況是，受訪者可能根本沒留意題目在問什麼，或在答題時敷衍以對（Oppenheimer et al. 2009; Jones et al. 2015）。

如果受訪者在填答網路調查問卷的過程中不夠專注，可能會造成測量誤差（*measurement error*）或題項無回應（*item nonresponse*），而影響到資料品質（Gummer et al. 2021）。也因為如此，如何偵測網路調查受訪者的專注程度，是網路調查方法研究上的重要議題。研究者指出，在網路自填問卷中加入注意力檢測題（*attention checks*），² 可用以評估受訪者在填答問卷時的專注程度。在注意力檢測題的設計上，一般的網路調查是在問卷中加上額外的題項，指示受訪者對該題須如何作答，藉由受訪者對該題的作答結果，評估受訪者是否通過注意力檢測。

在網路調查中，注意力檢測已使用地相當廣泛（Berinsky et al. 2014, 2016; Gummer et al. 2021）。儘管注意力檢測已廣為應用在網路調查中，嚴謹的方法研究仍不多見。但值得注意的是，在國際頂尖的調查方法、心理、政治、市場研究等領域的期刊中，近期出現多篇注意力檢測方面的方法研究（Alvarez et al. 2019; Anduiza and Galais 2016; Clifford and Jerit 2015; Curran 2016; Gummer et al. 2021; Kung et al. 2018;

1 最低滿意程度的概念，最早是由心理學者 Simon（1956）提出。

2 亦稱為篩選題（*screeners*）、陷阱題（*trap questions*）或紅鯊魚題（*red herring questions*）。

Liu and Wronski 2018; Silber et al. 2019)。相關方法研究的討論，多集中於如何設計注意力檢測題、注意力檢測題通過與否能否反映問卷資料品質、注意力檢測結果與其他資料品質指標間的關係（Berinsky et al. 2014; Jones et al. 2015; Curran 2016）。

然而，對於注意力檢測題是否會影響受訪者後續的作答行為，及是否會影響受訪者退出問卷的可能性，雖有學者點出這些問題的重要性，相關的方法研究仍極為罕見（Gummer et al. 2021）。由於網路自填問卷是一個網頁接一個網頁進行的，在注意力檢測題出現後，受訪者的作答行為、完訪意願是否受到影響，實為重要的研究課題。這項研究採用機率樣本（probability sample）的實驗性網路調查，³藉由將樣本隨機分派到設有注意力檢測題、未設注意力檢測題的問卷，分析注意力檢測題對受訪者退出問卷機率的影響，及後續題項作答行為的影響。在下節中，將討論相關文獻。第三節中，將說明本研究的資料蒐集方式、實驗設計、分析模型，以及所使用的變項。第四節將說明分析結果，末節則為結論。

二、相關文獻回顧與討論

網路調查是由受訪者自行填答問卷，不同於面訪、電訪由訪員進行問卷訪談，參與網路調查的受訪者在閱讀問卷題目、思考與填寫答案時可能會因為不夠專注，而造成測量誤差或題項無回應等情況，影

3 關於機率樣本、非機率樣本的網路調查，可參考 Yeager et al. (2011) 的討論。相較於自網路招募受訪者等非機率樣本的網路調查，機率樣本網路調查的母體有明確的範圍，而母體裡的單位有已知的中選機率，在由完訪資料進行推估時，不會有非機率網路調查母體不明的問題。

響到問卷資料品質 (Alvarez et al. 2019; Anduiza and Galais 2016)。在應用完訪資料進行推估或分析時，可能會因此產生估計上的偏誤，並影響到假設檢定的精確度及正確性。因此，在進行網路調查時，如何評估參與者是否專注、是否可能借助某些工具讓參與者在填答問卷時更為專注，是調查方法研究上的重要課題。

近年在調查方法、政治、心理研究等領域，對如何偵測網路調查參與者是否專注，是網路調查方法研究的重要議題 (Alvarez et al. 2019; Anduiza and Galais 2016; Clifford and Jerit 2015; Curran 2016; Gummer et al. 2021; Kung et al. 2018; Liu and Wronski 2018; Silber et al. 2019)。在網路調查中，常用來偵測參與者專注程度的方式，是在問卷中加入注意力檢測題，藉由完訪資料瞭解參與者在填答該題時是否有依照題目的指示作答，以評估受訪者在進行問卷填答時是否專注 (Alvarez et al. 2019)。常見的注意力檢測題設計，包含指引操弄檢測 (instructional manipulation checks, IMC)、指示應答題 (instructed-response items, IRI)、虛題 (bogus items) 等幾種類型 (Meade and Craig 2012; Gummer et al. 2021)。

指引操弄檢測要求受訪者在讀到題目時忽略該題的選項，依題目的指示進行操作 (例如，點選螢幕上的特定元件)，藉以瞭解受訪者是否有讀取題目的指示。指示應答題是指引操弄檢測的特殊題型，藉由在題組中加入額外的題目，請受訪者依該題的指示選擇特定選項 (例如，選擇「完全不同意」)，以瞭解受訪者是否讀到該題的指示。虛題是對有正確答案的題目 (例如，「水是濕的」)，請受訪者點選特定答案 (例如，選擇「不同意」)。⁴ 對注意力檢測題的設計，以下用兩個

4 參見 Gummer et al. (2021) 的討論。

例子加以說明。第一個例子來自 Kung et al. (2018: 270) 的單選題。

For this question, please select number six to demonstrate your attention.

第二個例子，來自 Berinsky et al. (2014: 740) (見圖 1)。這題是複選題，請受訪者忽略題目，選擇“ABC News”、“The Drudge Report”兩個選項。

儘管注意力檢測題已普遍應用在許多領域的網路調查中，對注意力檢測題的方法研究仍不多見 (Gummer et al. 2021)。以注意力檢測題與資料品質的關係而言，Oppenheimer et al. (2009) 指出，在網路自填問卷中加入指引操弄檢測，可偵測出哪些參與者沒有依照指示填答，有效找出不夠專注的參與者。相關研究發現，未通過注意力檢測題的受訪者對問卷的答題速度較快，且題項無回應較多，在題組方面較容易出現直線作答 (straight lining) 與答案不一致的情況，在開放題的

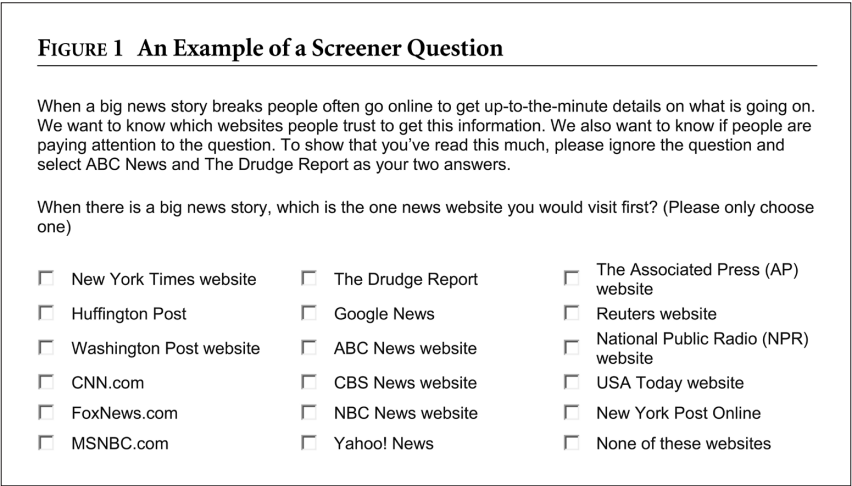


圖 1 注意力檢測題範例 (取自 Berinsky et al. 2014: 740, Figure 1)

填答字數較少 (Alvarez et al. 2019; Anduiza and Galais 2016; Atkeson et al. 2018; Gummer et al. 2021)。這些研究發現，注意力檢測題的檢測結果與其他資料品質指標間具有密切的關係。

有關注意力檢測題對於退出問卷的可能性、後續題項答題行為的影響，研究較為罕見。在 Berinsky et al. (2014) 操作的一項調查實驗中，將樣本隨機分派給兩份問卷，其中一份問卷將注意力檢測題設在研究者關心的敏感題組之前，另一份問卷則將注意力檢測題組置於問卷最末，以比較兩群樣本退出問卷的可能性、對敏感題組的答案分配是否有顯著差異。其分析結果發現，無論是退出問卷的比例或後續敏感題項的答案分配，兩群樣本均無顯著差異。Hauser and Schwarz (2015) 發現，在注意力檢測題出現之後，對複雜的推理測驗題，受訪者的答題表現較之前來得好。不同於 Hauser and Schwarz (2015)，Gummer et al. (2021) 發現注意力檢測題對後續的答題行為並無顯著影響，與 Berinsky et al. (2014) 的結果相近。在 Oppenheimer et al. (2009) 的研究中，針對未通過注意力檢測題的受訪者持續進行指引操弄檢測，直到通過檢測為止。Oppenheimer et al. (2009) 發現，對未通過注意力檢測題的受訪者重新進行指引操弄檢測，可提升受訪者的專注程度，進而提升資料品質。不同於 Oppenheimer et al. (2009)，Berinsky et al. (2016) 發現重複進行注意力檢測對問卷資料品質並無顯著影響，但會讓受訪者退出問卷的可能性變高。

有關注意力檢測題對後續作答行為可能存在的負面效應，研究者指出，認真填答網路自填問卷的受訪者在遇到注意力檢測題時，可能會覺得不被資料蒐集者信任，而對繼續填答問卷的可能性與後續作答行為產生不利影響 (Oppenheimer et al. 2009: 871)。研究者亦指出，在網路自填問卷中出現注意力檢測題這類警示，可能會讓受訪者覺得是

在監控的環境中作答，而產生社會讚許偏誤的問題（Clifford and Jerit 2015），或影響繼續填答問卷的意願（Berinsky et al. 2014: 744）。

自相關文獻的討論可以得知，對於注意力檢測題結果與其他資料品質指標的關聯，多數量化研究得到一致的結果，發現注意力檢測題的檢測結果與其他資料品質指標間有密切關係，表示注意力檢測題通過與否可作為個案的問卷資料品質指標。不過，在存在其他資料品質指標（如答題時間、題項無回應）的情況下，是否有必要採用注意力檢測題作為問卷資料品質的測量，文獻仍有不同的觀點。自前述文獻討論可知，受訪者有可能因為看到注意力檢測題而覺得不安或感到不受信任，以致影響後續的答題行為，甚或退出問卷填答。有關注意力檢測題對受訪者後續答題行為、完成問卷與否的影響效果，實有必要做更深入的研究。

有關注意力檢測題對於同一份問卷後續答題行為的影響，現有文獻幾乎都是自「問卷」層級（level）觀察注意力檢測題結果與整份問卷資料品質的關係，如整份問卷的填答時間、無回應題數、對開放題的填答情況、對敏感題項的填答情形等。然而，不同於紙本問卷可一次看到所有題項，網路自填問卷在進行填寫時，題項是逐頁出現的。亦即，完成某個網頁的題項後，必須在切換至下一個頁面時，才能看到下一個網頁的題項。因此，自「題項」層級切入，分析注意力檢測題對後續題項作答行為的影響，有其學理上的重要性。

基於前述，本文以中央研究院人文社會科學研究中心轄下調查研究專題中心（以下簡稱調研中心）建置的「網路調查會員資料庫」中的會員為對象，進行網路調查實驗，以探討注意力檢測題對後續題項作答行為、退出問卷機率的影響。在後續題項作答行為方面，是採用個人、題項的兩層級模型，針對後續題項的回應時間（response latency）

進行分析，以納入題項層級的影響效果。以題項回應時間作為應變項的主要原因是，題項回應時間可反映受訪者在答題過程中投入的認知努力（cognitive efforts），如果在答題上投入的努力愈多或處理的資訊愈多，題項回應時間會愈長（Bassili and Scott 1996; Couper and Kreuter 2013; Leiner 2019; Turner et al. 2015）。就作者所知，在注意力檢測題對後續題項作答行為的研究中，本文是唯一考慮到題項層級，並以題項回應時間作為應變項的研究。

在注意力檢測題對退出問卷機率的影響方面，本文的研究設計不同於 Berinsky et al.（2014: 744）只比較兩組樣本（一組將注意力檢測題置於問卷中間，另一組則將注意力檢測題置於問卷最末）在完填問卷的比例是否存在顯著差異。⁵ 首先，本文同時納入完填問卷、部分填答問卷的樣本，並用到會員的個人基本資訊，在控制個人特質之下，分析注意力檢測題對受訪者完填問卷與否的影響。其次，在調查實驗分組中，我們設計了兩個注意力檢測題，以及問卷前、後兩個放置位置，以分析注意力檢測題的放置位置、題目設計是否影響後續作答行為與退出問卷的機率。

有關注意力檢測題對退出問卷機率、後續題項回應時間的影響，以下分別說明可能的影響效果。關於前者，如果受訪者因注意力檢測題而更專注於問卷填答，會有較高的動機或動力完成問卷；反之，如果受訪者對注意力檢測題感到不悅或不安，可能會選擇退出問卷而造成部分填答。至於注意力檢測題對退出問卷機率的影響效果，可以預先設想的是，如果受訪者會因注意力檢測題而更專注於問卷填答，在

5 Berinsky et al.（2014）將其中一組樣本的注意力檢測題置於問卷最末，由於其後已經沒有題目，無法比較此組樣本的后續作答行為是否與另一組樣本相同。

後續題項的作答上，速度可能會放緩，而使得題項回應時間增加。反之，如果受訪者因為對注意力檢測題感到不舒服，在後續題項的作答上，可能會較為草率，並加快填答速度。考量不同心理因素的作用，在後文中將透過實際的資料分析，檢視以上哪些（個）假設可得到實證結果的支持。

三、資料來源與研究方法

（一）資料來源

這項研究採用自行設計的問卷，以實驗設計的方式將調查對象分為五群，委託調研中心執行網路調查。調研中心自 2015 年開始建置「網路調查會員資料庫」，其會員並非自網路招募而來，而是來自調研中心執行的各類機率抽樣調查。這些調查是在訪問結束時，詢問受訪者是否有意願參與調研中心後續執行的網路調查。對有意願參與並提供電子郵件信箱或手機號碼者，再以電子郵件或手機簡訊邀請上網填答一份包含個人基本資訊在內的短卷，完成後即成為調研中心「網路調查會員資料庫」的會員。由於「網路調查會員資料庫」的會員為既有抽樣調查的完訪者，而調研中心執行的抽樣調查都是由嚴謹的隨機抽樣程序抽出樣本，據以建立的「網路調查會員資料庫」具有機率樣本的特性。

本研究以「網路調查會員資料庫」的會員為對象，委託調研中心執行網路調查，調查執行期間為 2021 年 4 月 19 日至 5 月 3 日。⁶ 針

6 在執行調查之前，調查計畫有先取得中央研究院人文與社會科學研究倫理委員會的審查通過證明。

對未填答者，期間最多進行三次催收，催收日期分別為 4 月 21、24、29 日。完成問卷填答的會員，可得到由計畫提供的 50 元超商現金抵用券。由調研中心寄發的網路調查邀請函，計有 7,516 封，其中成功寄發的有 7,400 封。最終的填答問卷數共計 3,875 案，其中部分填答 105 案，完成填答的則有 3,770 案，完成率為 50.95%。⁷

在問卷主題方面，是以個人對調查、資料的態度作為主要議題，包含對不同調查機構的信任程度、對民意調查的態度、對各類個人資訊的敏感程度、接受調查與否的考量因素、是否願意讓特定行政登記資料作為學術研究之用、是否同意蒐集個人特定資訊（如臉書、基因檢測資料）資料作為學術研究之用、對政府公務調查與資料的態度等。除這些題項外，也納入基本人口特徵、參與網路調查的經驗等。如果不將注意力檢測題計入，總計有 61 個問項。其中的 52 個問項，是無法跳答的題項。

在實驗設計上，是由調研中心事先將調查對象隨機分為五組。其中一組為對照組（control group），不設注意力檢測題。考量注意力檢測題的題目、位置可能影響受訪者的作答與退出行為，我們設計了兩個注意力檢測題，並選擇了兩個位置，因而共有 4 個（2×2）實驗組。在注意力檢測題的設計上，參酌 Kung et al.（2018）的設計，在題目中加入「以顯示您的閱讀專注力」。五組樣本的問卷，除注意力檢測題外，其餘題項完全相同。兩個注意力檢測題，包括以下的 A、B 兩題：

A. 這題請您選擇最中間的答案，以顯示您的閱讀專注力

B. 有人說：人口統計對國家的人口政策完全不重要，請您選擇非常同意，以顯示您的閱讀專注力

7 在完成率的計算上，是以完成填答數為分子，寄發成功數為分母。

這兩題的答案，同樣有後列五個選項：(1) 非常不同意；(2) 不同意；(3) 無所謂同不同意；(4) 同意；(5) 非常同意。注意力檢測題放置的位置有二，其一是位於問卷第 17、18 題之間（問卷前半部），另一則是位於第 53、54 題之間（問卷後半部）。四個（2×2）實驗組的代號分別如後：A1（A 注意力檢測題置於問卷前半部）、B1（B 注意力檢測題置於問卷前半部）、A2（A 注意力檢測題置於問卷後半部）、B2（B 注意力檢測題置於問卷後半部）。對照組（未放置注意力檢測題）則以 C 代表。

將注意力檢測題安排於第 17 題及第 53 題之後，一方面是因為這兩個位置分別位於問卷前半部、後半部，另一方面則是因為兩個位置分別位於兩組態度題組的中間。既有方法研究指出，在問卷包含態度題組或量表的情況下，由於題組有著相同的答題設計，可能使得受訪者在填答題組時不仔細閱讀題目，而產生敷衍作答的行為，因而主張將注意力檢測題置於題組中間（Alvarez et al. 2019; Kung et al. 2018）。在本文的注意力檢測題設計上，也依循這樣的準則，將兩個注意力檢測題分別置於兩組態度題組的中間。

除了問卷資料外，研究中也用到「網路調查會員資料庫」的會員基本資料，包括性別、年齡、教育程度、招募途徑、歷次網調參與情況等。其中的招募途徑，係指透過面訪、市話電訪、手機電訪或手機簡訊網調加入會員。⁸ 歷次網調參與情況則指歷次會員網路調查的參與次數。除前述周邊資料（paradata）外，運用網路調查系統自動記錄

8 調研中心手機簡訊網調的抽樣方式與手機電訪調查一樣，採分層兩階段等距抽樣，分層依據手機門號千門碼（前 6 碼）的使用率高低分群，每一分層中第一階段以等距抽樣抽出中選門號的千門碼，第二階段將中選的手機門號採末三碼隨機抽出中選樣本的手機號碼，以寄送簡訊的方式邀請手機持有人上網填答問卷。

的每個題項開始、結束時間，可計算參與這次網路調查的會員停駐在每個題項上的時間；另可瞭解中途退出問卷填答者，是在哪個題項退出問卷。⁹問卷、周邊資料兩者的整合資料，對這項研究提供了相當充足的資訊。

（二）分析方法

藉由前述網路調查實驗的問卷資料、周邊資料，除了針對五組樣本進行敘述統計分析外，將分別採用適切的分析模型，分析注意力檢測題的設定對退出問卷與否、題項作答時間兩類應變項的影響效果。以下分別說明使用的分析模型。

1. 退出問卷與否的模型

對於個人是否退出問卷填答，本研究採用兩種模型進行分析，其一是 Cox 比例風險模型（Cox proportional hazard model），¹⁰ 另一為 logit 模型。對具有「個人—時間」兩個層級的資料，既有研究常會採用 Cox 比例風險模型或 logit 模型，分析個人自某個狀態轉換為另一狀態的可能性（Ingram and Kleinman 1989）。由於本研究使用的資料有「個人—題項」兩個層級，也因應使用 Cox 比例風險模型、logit 模型，分析個人自問卷填答轉為退出問卷的可能性。然而，如同既有文獻的說明，兩種模型的分析結果在解讀上是有差異的（如 Ingram and Kleinman 1989）。在本文的架構下，Cox 比例風險模型觀察的是個人退出問卷相對於繼續填答問卷的相對風險（relative risk），或稱風險比

9 調研中心所使用的電腦輔助調查（computer-assisted interviewing, CAI）系統，是由該單位自行開發，整合了面訪、電訪及網路調查等不同模式為單一系統。本研究在相關周邊資料的取得與使用上，有經 IRB 審查通過。

10 可參考 Cox and Oakes（1984）。

(hazard ratio)；自 logit 模型則可得到個人退出問卷、繼續填答問卷機率的比值，或稱勝算比 (odds ratio)。

以下分別針對兩類模型的設定做說明。在 Cox 比例風險模型的設定方面，是分析受訪者隨著題項 (t) 的進行，退出問卷填答的可能性。假設退出問卷的風險函數，可用下式表示。

$$h(t) = h_0(t) \cdot \exp(\beta_{A1} \cdot A1 + \beta_{B1} \cdot B1 + \beta_{A2} \cdot A2 + \beta_{B2} \cdot B2 + X \cdot \beta_X) \quad (1)$$

其中， $h_0(t)$ 代表無法直接估計的基線風險 (baseline hazard)， X 包含模型中的控制變項， β_{A1} 、 β_{B1} 、 β_{A2} 、 β_{B2} 為 $A1$ 、 $B1$ 、 $A2$ 、 $B2$ 四個實驗組 (參照組為對照組 C) 對應的參數， β_X 則為 X 對應的參數。

在 logit 模型方面，則以進入問卷填答的受訪者是否中途退出問卷作為應變項。若中途退出問卷，則設虛擬變項值為 1，否則為 0。在控制前述的 X 變項之下，觀察 $A1$ 、 $B1$ 、 $A2$ 、 $B2$ 各實驗組相對於對照組在退出問卷的機率上是否有顯著差異。對於 logit 模型的方程式，在此就不贅述。不論前述的 Cox 比例風險模型或 logit 模型，在分析上並非以完訪問卷資料作為分析資料，而是以「所有進入問卷填答」（不論是否完成整份問卷的填答）的問卷資料作為資料來源。也因為如此，在控制變項的選用上，主要是以個人加入「網路調查會員資料庫」時所提供的資訊為主，包括個人的性別、出生年次、教育程度等。其中的出生年次，在實際分析時，將該變項轉換為 2021 年接受調查時的年齡，並加入年齡平方項以控制可能的非線性效果。教育程度則依個人加入會員時的教育程度區分為四類：高中職或以下、專科、大學、研究所。除了這些人口特徵變項外，也納入個人過去參與「網路調查會員資料庫」的次數，作為控制變項之一。

自第 (1) 式實驗組樣本對應係數 (β_{A1} 、 β_{B1} 、 β_{A2} 、 β_{B2}) 的估計

值，可得知實驗組受訪者退出問卷的機率，是否與未設注意力檢測題的對照組存在顯著差異。從 logit 模型的分析結果，亦可判斷實驗組、對照組樣本在中途退出問卷的機率上是否有顯著不同。

2. 題項回應時間模型

為瞭解不同注意力檢測題的設定是否影響個人的題項回應時間，在此將資料分為個人、題項兩個層級，以分析位於注意力檢測題之後的題項，在題項回應時間上是否顯著較長或較短，抑或沒有顯著差異。在這部分的分析中，是以無法跳答的 52 個題項作為分析對象，而非所有的 61 個題項。主要的考量是，對可能跳答的題項，在題項回應時間上會有部分樣本存在缺漏值。但為便於說明，以下的題項回應時間模型是以所觀察的題項位於 61 個題項中的順位表示題項所在位置。對於參與問卷填答的受訪者（不論是否完成全部的題項填答），針對第 i 位受訪者花在第 j 題的題項回應時間（ Y_{ij} ），設定如下的兩層級模型。

$$\begin{aligned}
 g(Y_{ij}) = & \alpha_0 + \alpha_{A1} \cdot A1_i + \alpha_{B1} \cdot B1_i + \alpha_{A2} \cdot A2_i + \alpha_{B2} \cdot B2_i \\
 & + \alpha_{I1} \cdot I(Item_{ij} > 17) + \alpha_{I2} \cdot I(Item_{ij} > 53) \\
 & + \alpha_{A1I1} \cdot A1_i \cdot I(Item_{ij} > 17) + \alpha_{B1I1} \cdot B1_i \cdot I(Item_{ij} > 17) \\
 & + \alpha_{A2I2} \cdot A2_i \cdot I(Item_{ij} > 53) + \alpha_{B2I2} \cdot B2_i \cdot I(Item_{ij} > 53) \\
 & + \mathbf{X}_i \cdot \boldsymbol{\alpha}_X + \mathbf{Z}_j \cdot \boldsymbol{\alpha}_Z + u_i + \epsilon_{ij}
 \end{aligned} \tag{2}$$

其中， $g(\cdot)$ 為連結函數（link function），由於題項回應時間（ Y_{ij} ）為右偏分配，在分析中是將題項回應時間取自然對數後的數值作為應變項。 $A1$ 、 $B1$ 、 $A2$ 、 $B2$ 變項的意義與前文相同。 $I(Item_{ij} > 17)$ 、 $I(Item_{ij} > 53)$ 兩個虛擬變項，分別代表題項位置是否位於注意力檢測題所在位置的 17、53 題之後。 $A1_i \cdot I(Item_{ij} > 17)$ 、 $B1_i \cdot I(Item_{ij} > 17)$ 、 $A2_i \cdot I(Item_{ij} > 53)$ 、 $B2_i \cdot I(Item_{ij} > 53)$ 分別為注意力檢測題組與題項位置虛擬變項的交叉項。

X_i 、 Z_j 則分別代表受訪者、題項兩個層級的控制變項。 u_i 代表個人層級的隨機效果 (random effects)，假設 $u_i \sim N(0, \sigma_u^2)$ 。 ϵ_{ij} 為模型誤差項 (error term)，假設 $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$ 。後文採用多層級混合效果一般化線性模型 (multilevel mixed-effects generalized linear model) 作為分析模型。

在題項回應時間的測量方面，是參考既有文獻的作法 (Gummer et al. 2021; Olson and Smyth 2015)。由於調查系統中，會記錄每題的作答開始時間、提交時間，自相關紀錄，可計算每個題項的回應時間 (以秒為單位)。對模型的控制變項，參酌相關研究 (Olson and Smyth 2015; Couper and Kreuter 2013)，除第 (1) 式考慮的個人層級控制變項外，也納入題項特性作為題項層級的控制變項。在題項特性方面，基於受訪者在進行不同類型題項的作答時，在作答上花了解讀題意、萃取答案、選擇答案的時間可能不同，將題項類型取分為三種類型：事實題 (factual question)、態度題 (attitude question)，及其他主觀題。

在本研究中，樣本是否適用於設有注意力檢測題的問卷出於隨機分派，而實驗組、對照組除注意力檢測題外，其他題項的內容、序次完全一致。自第 (2) 式交叉變項對應係數 (α_{A1I1} 、 α_{B1I1} 、 α_{A2I2} 、 α_{B2I2}) 的估計值，可得知在設有注意力檢測題的問卷中，注意力檢測題對於其後題項的回應時間是否有顯著影響。

四、分析結果

在本節中，將先呈現各變項的敘述統計概況，接著分別呈現退出問卷、題項回應時間模型的分析結果。

（一）敘述統計結果

對文中使用的變項，將樣本依實驗組、對照組分群，將變項的平均值列於表 1。在表 1 中，對連續變項，亦於括號內列出對應的標準差。自表 1 可知，兩性的比例大致相當，男性約占 48.46%。各實驗組、對照組樣本的男性比例，介於 45.86% 至 51.31% 之間；平均年齡則為 41 至 43 歲之間。教育程度方面，以大學的占比最高，約占半數左右，其次為研究所（約占 24%），而專科、高中職或以下兩者的比例大致相當（約各占 13%）。這可能是因為「網路調查會員資料庫」主要透過電子郵件信箱、手機號碼邀請個人加入會員，而使得教育程度偏高。至於會員的來源，則以手機簡訊網調為主（約為四成），其次為市話電訪（約占 34%）、面訪（約占 21%），而以手機電訪最少（約占 3%）。過去參與「網路調查會員資料庫」網路調查的平均次數，約為 4 次。

對於部分填答（或退出問卷）此一應變項，自表 1 可知，約有 3% 左右的受訪者在進入問卷填寫後退出。至於實驗組、對照組分群樣本，退出問卷的比例則介於 2.65% 至 3.01% 之間，然而彼此間的差異並不大。為觀察受訪者退出問卷的機率隨題項的變化情況，針對實驗組與對照組五組樣本，分別繪製退出問卷的存活函數估計值如圖 2。為便瞭解注意力檢測題在圖形中所在的大概位置，在第 17、53 題處以垂直線標示。¹¹ 自圖 2 可以看出，對不同的樣本群而言，退出問卷機率隨題項的變化大致是相似的。在圖 3 中，以另一種方式呈現五群樣本的退出問卷比例隨題項的變化。以「無檢測題」的對照組樣本為例，該圖形顯示

11 在 Stata 的實地操作上，所撰寫的語法是：sts graph, survival by (group)。其中的“group”代表樣本所屬群組（對照組、四個實驗組共 5 組樣本）變項。

表 1 變項基本統計特性：依實驗組與對照組分群

	未設注意力檢測的群組	設有注意力檢測的群組				所有樣本
		題 A 置於問卷 前半部	題 B 置於問卷 前半部	題 A 置於問卷 後半部	題 B 置於問卷 後半部	
男性	47.55%	47.70%	51.31%	45.86%	50.00%	48.46%
年齡	42.76 (13.06)	41.39 (12.80)	41.59 (11.96)	42.01 (12.62)	42.02 (13.04)	41.94 (12.71)
教育程度						
高中職或以下	14.04%	14.16%	14.70%	11.65%	13.02%	13.50%
專科	12.72%	12.76%	12.07%	14.29%	13.40%	13.06%
大學	50.07%	49.62%	48.95%	50.25%	50.00%	49.78%
研究所	23.18%	23.47%	24.28%	23.81%	23.58%	23.66%
招募管道						
面訪	20.93%	21.05%	19.16%	23.10%	21.91%	21.24%
市話電訪	33.38%	33.55%	34.78%	33.46%	35.05%	34.04%
手機電訪	3.34%	2.55%	2.76%	4.89%	2.71%	3.02%
手機簡訊網調	43.58%	42.86%	43.31%	38.60%	40.34%	41.70%
會員網調參與次數	3.87 (4.00)	3.64 (3.87)	3.73 (3.94)	3.97 (4.03)	4.10 (4.05)	3.86 (3.98)
部分填答	2.65%	2.76%	2.76%	3.01%	2.71%	2.71%
題項回應時間（原始資料）						
事實題	9.60 (22.91)	9.37 (22.62)	9.13 (16.02)	9.58 (22.10)	9.49 (21.13)	9.43 (21.13)
態度題	10.34 (24.20)	10.18 (20.93)	9.97 (20.76)	10.14 (25.18)	10.65 (24.21)	10.26 (23.14)
其他主觀題	7.29 (15.50)	7.12 (16.59)	7.01 (15.13)	7.05 (15.09)	7.67 (20.59)	7.23 (16.71)

表 1 變項基本統計特性：依實驗組與對照組分群（續）

	未設注意 力檢測的 群組	設有注意力檢測的群組				所有 樣本
		題 A	題 B	題 A	題 B	
		置於問卷 前半部	置於問卷 前半部	置於問卷 後半部	置於問卷 後半部	
題項回應時間（99%）						
事實題	8.79 (9.58)	8.65 (9.76)	8.66 (9.60)	8.79 (10.54)	8.74 (10.20)	8.72 (9.95)
態度題	9.40 (7.99)	9.32 (8.08)	9.13 (7.76)	9.23 (7.87)	9.62 (8.26)	9.34 (8.00)
其他主觀題	6.72 (5.91)	6.62 (5.84)	6.51 (5.56)	6.49 (5.65)	6.81 (6.10)	6.63 (5.82)
題項回應時間（98%）						
事實題	8.55 (7.84)	8.40 (7.87)	8.42 (7.76)	8.48 (8.16)	8.45 (8.03)	8.46 (7.94)
態度題	9.21 (7.10)	9.10 (7.04)	8.93 (6.79)	9.02 (6.89)	9.38 (7.26)	9.13 (7.02)
其他主觀題	6.57 (5.19)	6.46 (5.11)	6.38 (4.91)	6.34 (4.95)	6.62 (5.31)	6.47 (5.10)
題項回應時間（95%）						
事實題	8.19 (6.53)	8.03 (6.38)	8.05 (6.40)	8.08 (6.55)	8.07 (6.57)	8.08 (6.49)
態度題	8.81 (5.86)	8.70 (5.73)	8.59 (5.67)	8.63 (5.64)	8.96 (5.96)	8.74 (5.77)
其他主觀題	6.31 (4.41)	6.20 (4.28)	6.16 (4.20)	6.11 (4.22)	6.32 (4.45)	6.22 (4.31)
受訪者人數	755	784	762	798	776	3,875

註：1. 平均值列於表中，括號內則為標準差。

2. 題項回應時間的統計數據是由個人—題項層級資料計算，其餘數據由個人層級資料計算。

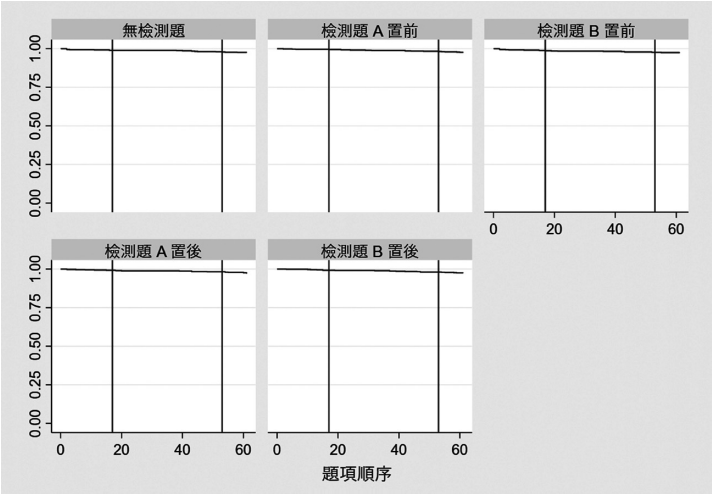


圖 2 Kaplan-Meier 退出填答的存活函數估計值：
依實驗組與對照組分群

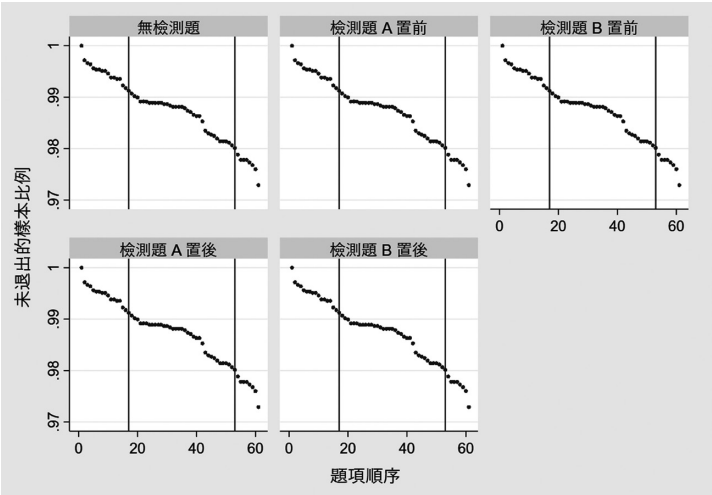


圖 3 未退出問卷的樣本比例：依實驗組與對照組分群

各題填答的樣本數占該群原有樣本數的比例。自圖 3 的五個圖形可粗略得知，五群樣本的趨勢大致相若。儘管如此，五群樣本在退出問卷的機率上是否有顯著差異，仍需仰賴較嚴謹的分析解答（見下一小節）。

題項回應時間是根據網路調查系統的自動記錄，以進入題項至提交答案兩個時點的時間差異作為測量，並以秒為單位。在表 1 中，依照題項類型，分別列出事實題、態度題、其他主觀題的平均題項回應時間。自調查系統產出的原始數據可以得知，不論是哪類題型，題項回應時間的標準差都偏高。而在實際檢視各題回應時間的分配時亦發現，各題的回應時間都存在數值很大的離群值（outlier）。可能的原因是，受訪者在網路問卷作答的過程中，可能因為要處理其他事情或有其他事情插進來，而延緩提交答案。

既有的題項回應時間研究，多未處理離群值的問題，而是以原始資料進行分析（Read et al. 2021）。對離群值的處理，由於既有文獻未提供可依循的準則，作者對題項回應時間超過某一臨界值者，將題項回應時間重新編碼為臨界值。以表 1 所列的「題項回應時間（95%）」為例，是逐題設算各題回應時間分配位於 95% 的臨界值為多少，對該題回應時間高於此一臨界值者，以臨界值取代原始的題項回應時間。「題項回應時間（98%）」、「題項回應時間（99%）」亦可類推。在將原始的題項回應時間做了前述處理後，由表 1 可知，不論哪類題型，題項回應時間的平均值、標準差都較未處理的情況下為低，顯示離群值所造成的干擾下降。除前述處理方式之外，對於偏高的題項回應時間，亦嘗試以其他三種方式處理，包含將該題偏高的數值視為遺漏值（missing value），自分析資料中刪除，或是以該題回應時間的平均值或中位數取代。

自表 1 可知，對於題項回應時間高於 95%、98%、99% 的數值，

如以 95%、98%、99% 臨界值取代，在事實題、態度題、其他主觀題三類題型中，以態度題的平均題項回應時間最長，其次為事實題，最短的則為其他主觀題。以未經處理的原始資料為例，對態度題、事實題、其他主觀題，所有樣本的平均題項回應時間分別為 10.26、9.43、7.23 秒。在經過臨界值重新編碼的情況下，若臨界值對應的百分比愈低（如 95%），由於有較多的題項回應時間會被重新編碼，題項回應時間的平均值、標準差均會較低。以「題項回應時間（95%）」為例，對態度題、事實題、其他主觀題，所有樣本的平均題項回應時間分別為 8.74、8.08、6.22 秒。¹²

為便讀者瞭解題項回應時間的分配情況，在此以第 2 題「請問您是民國哪一年出生的」為例，將該題未經處理的題項回應時間分配列於圖 4-1；經過 99%、98%、95% 臨界值處理的題項回應時間分配，則分別列於圖 4-2、圖 4-3、圖 4-4。自圖 4-1 可知，有部分樣本的題項回應時間明顯偏高，最高值為 408 秒；在經過 99%、98%、95% 臨界值處理後，題項回應時間的最高值分別為 44、29、17 秒。其他的題項，在未經處理的題項回應時間方面，也都有類似第 2 題的分配。自圖 4-1、圖 4-2、圖 4-3、圖 4-4 可知，無論題項回應時間資料是否經過處理，或以哪個臨界值處理，均呈右偏分配。至於將偏高的題項回應時間視為遺漏值、以平均值取代，以中位數取代三種資料處理方式，經處理後的題項回應時間也均呈右偏分配。¹³ 基於前述說明，且

12 考慮篇幅，對於偏高的題項回應時間視為遺漏值、以平均值取代、以中位數取代三種資料處理方式，在表 1 中並未列出經處理後的題項回應時間平均值、標準差。有興趣的讀者，可向作者索取相關數據。

13 基於篇幅的考量，對於第 2 題回應時間偏高的數值視為遺漏值、以平均值取代，以中位數取代三種資料處理方式，在文中並未呈現其處理後的題項時間分配。有興趣的讀者，可向作者索取統計圖表。

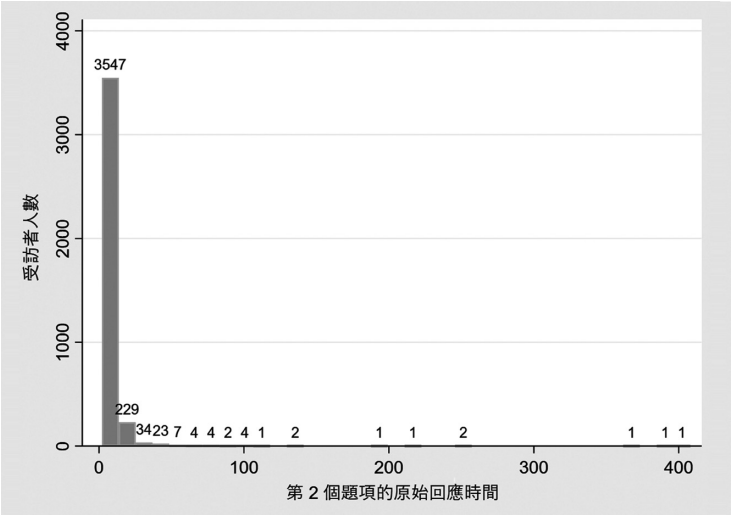


圖 4-1 第 2 個題項未經處理的題項回應時間分配

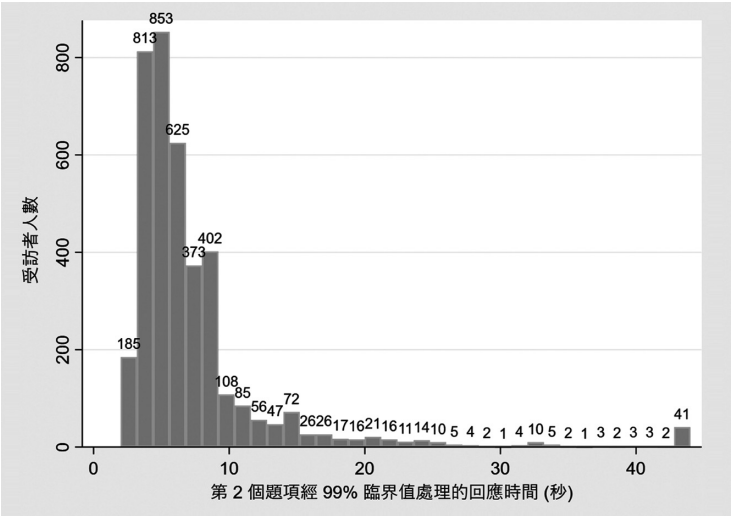


圖 4-2 第 2 個題項經 99% 臨界值處理的題項回應時間分配

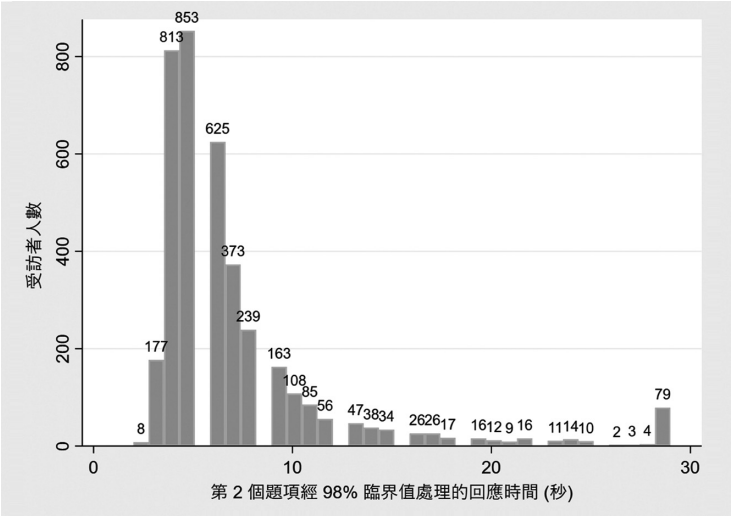


圖 4-3 第 2 個題項經 98% 臨界值處理的題項回應時間分配

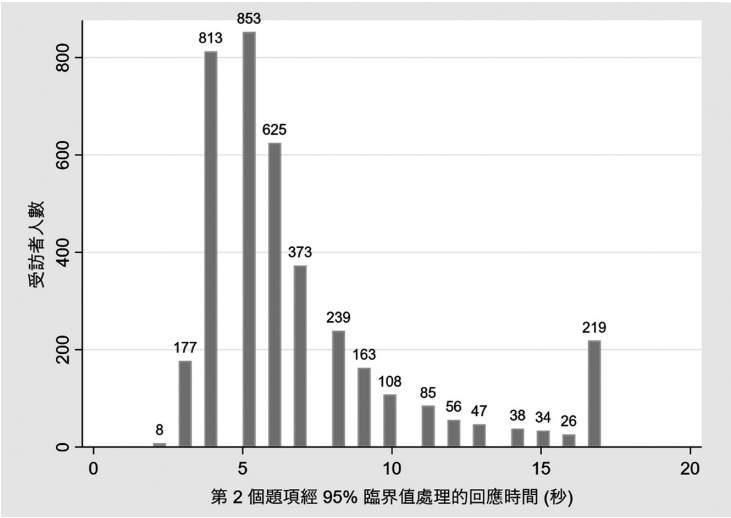


圖 4-4 第 2 個題項經 95% 臨界值處理的題項回應時間分配

題項回應時間相關方法研究多以取自然對數後的數值作為應變項（如 Callegaro et al. 2009; Johnson 2004），本文亦將經處理後的題項回應時間再取自然對數，以作為應變項。

（二）退出問卷與否模型的結果

對於受訪者退出問卷或部分填答與否，表 2 分別列出 Cox 比例風險模型、logit 模型的分析結果。表中 Cox 比例風險模型、logit 模型顯示的數值，分別為風險比、勝算比估計值。Cox 比例風險模型旨在分析隨著題項的進展，受訪者退出問卷的機率。若某個變項的風險比估計值大於 1，表示該變項對退出問卷的風險具正向影響；若小於 1，則表示有負向影響。不同於 Cox 比例風險模型，logit 模型意在分析受訪者部分填答與否。若某個變項的勝算比估計值大於 1，代表該變項對部分填答的機率具正向影響；若小於 1，則代表有負向影響。雖然兩者的概念不同，表 2 顯示，各實驗組與對照組間並無顯著差異。顯示受訪者退出問卷（或部分填答）的機率，不會受到注意力檢測題設定

表 2 退出問卷與否模型

	(1) Cox 比例風險模型	(2) Logit 模型
注意力檢測群組（ref=對照組）		
A1	0.931 (0.297)	0.929 (0.303)
A2	1.115 (0.341)	1.122 (0.352)
B1	1.201 (0.362)	1.212 (0.376)
B2	1.083 (0.335)	1.095 (0.348)

表 2 退出問卷與否模型（續）

	(1) Cox 比例風險模型	(2) Logit 模型
男性	0.695* (0.139)	0.692* (0.142)
年齡	0.931* (0.037)	0.931* (0.038)
年齡平方項	1.001** (0.000)	1.001** (0.000)
教育程度（ref=高中職或以下）		
專科	0.759 (0.237)	0.758 (0.246)
大學	0.590** (0.154)	0.584** (0.157)
研究所	0.661 (0.196)	0.652 (0.198)
招募管道（ref=手機簡訊網調）		
面訪	1.078 (0.447)	1.070 (0.452)
市話電訪	1.123 (0.265)	1.119 (0.271)
手機電訪	0.781 (0.472)	0.778 (0.483)
會員網調參與次數	0.760*** (0.043)	0.758*** (0.044)
常數項		0.218 (0.214)
Log-likelihood	-824.2	-440.5
觀察值個數／受訪者人數	233,512	3,875

註：1. 風險比、勝算比分別呈現於兩個模型，括號內為標準誤。

2. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ 。

的影響。亦即，在設有注意力檢測題的情況下，不論是採用 *A* 或 *B* 題型，或者置於問卷前半或後半，均不會影響受訪者退出問卷（或部分填答）的可能性。

至於控制變項，儘管 Cox 比例風險模型、logit 模型估計結果的詮釋不同，自表 2 可知，兩個模型的分析結果是相近似的。男性、大學教育程度者，退出問卷的可能性相對較低。在年齡方面，退出問卷的可能性會先隨年齡的增加而降低，而後再提升，呈現 U 形的影響效果。另外，如果受訪者過去參與會員網路調查的次數愈多，退出問卷的可能性愈低。

（三）題項回應時間模型的結果

對題項回應時間模型，針對題項回應時間採用 95%、98%、99% 臨界值處理後的資料，以及未處理的原始資料分析後的結果，分別列於表 3 的第 2 至 5 欄。以未經處理的資料而言（最後一欄），在控制個人、題項因素之下，四個實驗組虛擬變項均無顯著效果。題項位於第 17 題之後、第 53 題之後（ $I(Item_{ij} > 17)$ 、 $I(Item_{ij} > 53)$ ）兩個虛擬變項，均呈顯著的正向影響，表示對位於第 17 題或第 53 題之後的題項，受訪者的題項回應時間較長。另外， $A2_i \cdot I(Item_{ij} > 53)$ 、 $B2_i \cdot I(Item_{ij} > 53)$ 兩個交叉項均呈正向顯著，表示若受訪者所屬的組別是將注意力檢測題（*A* 或 *B* 題）設在問卷後半部（位於第 53、54 題之間）的群組，對於位於該檢測題之後的題項，受訪者花在答題的時間會顯著增加。至於離群值經 95%、98%、99% 臨界值處理後的分析結果，無論是實驗組虛擬變項、題項位置虛擬變項（ $I(Item_{ij} > 17)$ 、 $I(Item_{ij} > 53)$ ）或是實驗組虛擬變項與題項位置虛擬變項的交叉項，影響效果均與未經處理的結果相近。

表 3 題項回應時間自然對數值的多層級混合效果一般化線性模型：
偏高數值以臨界值取代

	(1) 大於 95% 的 數值以 95% 臨界值取代	(2) 大於 98% 的 數值以 98% 臨界值取代	(3) 大於 99% 的 數值以 99% 臨界值取代	(4) 資料未經 處理 (原始資料)
注意力檢測群組 (ref=對照組)				
A1	0.008 (0.014)	0.007 (0.015)	0.007 (0.015)	0.005 (0.016)
A2	0.003 (0.015)	0.003 (0.015)	0.003 (0.015)	0.003 (0.016)
B1	-0.019 (0.014)	-0.018 (0.014)	-0.018 (0.015)	-0.018 (0.015)
B2	0.005 (0.014)	0.007 (0.014)	0.008 (0.015)	0.010 (0.015)
題項位置				
$I(Item > 17)$	0.250*** (0.005)	0.250*** (0.005)	0.249*** (0.005)	0.249*** (0.005)
$I(Item > 53)$	0.263*** (0.006)	0.260*** (0.007)	0.260*** (0.007)	0.259*** (0.007)
注意力檢測群組×題項位置				
$A1 \cdot I(Item > 17)$	0.007 (0.006)	0.009 (0.007)	0.010 (0.007)	0.011 (0.007)
$B1 \cdot I(Item > 17)$	-0.001 (0.006)	-0.001 (0.007)	-0.001 (0.007)	-0.002 (0.007)
$A2 \cdot I(Item > 53)$	0.086*** (0.010)	0.088*** (0.011)	0.088*** (0.011)	0.089*** (0.011)
$B2 \cdot I(Item > 53)$	0.067*** (0.010)	0.067*** (0.011)	0.067*** (0.011)	0.065*** (0.011)

**表 3 題項回應時間自然對數值的多層級混合效果一般化線性模型：
偏高數值以臨界值取代（續）**

	(1) 大於95%的 數值以95% 臨界值取代	(2) 大於98%的 數值以98% 臨界值取代	(3) 大於99%的 數值以99% 臨界值取代	(4) 資料未經 處理 (原始資料)
題項類型（ref=事實題）				
態度題	0.074*** (0.004)	0.074*** (0.004)	0.073*** (0.004)	0.074*** (0.004)
其他主觀題	-0.335*** (0.005)	-0.337*** (0.005)	-0.336*** (0.005)	-0.336*** (0.006)
男性	-0.000 (0.009)	-0.002 (0.009)	-0.002 (0.009)	-0.004 (0.010)
年齡	0.002 (0.002)	-0.000 (0.002)	-0.001 (0.002)	-0.001 (0.002)
年齡平方項	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
教育程度（ref=高中職或以下）				
專科	-0.064*** (0.017)	-0.071*** (0.018)	-0.073*** (0.018)	-0.074*** (0.018)
大學	-0.107*** (0.014)	-0.113*** (0.014)	-0.115*** (0.014)	-0.116*** (0.015)
研究所	-0.140*** (0.015)	-0.146*** (0.016)	-0.148*** (0.016)	-0.148*** (0.016)
招募管道（ref=手機簡訊網調）				
面訪	0.049*** (0.015)	0.051*** (0.016)	0.051*** (0.016)	0.051*** (0.016)
市話電訪	0.061*** (0.012)	0.062*** (0.012)	0.062*** (0.013)	0.062*** (0.013)
手機電訪	0.110*** (0.026)	0.120*** (0.027)	0.124*** (0.027)	0.127*** -0.002

表 3 題項回應時間自然對數值的多層級混合效果一般化線性模型：
偏高數值以臨界值取代（續）

	(1) 大於 95% 的 數值以 95% 臨界值取代	(2) 大於 98% 的 數值以 98% 臨界值取代	(3) 大於 99% 的 數值以 99% 臨界值取代	(4) 資料未經 處理 (原始資料)
會員網調參與次數	-0.002 (0.001)	-0.002 (0.002)	-0.002 (0.002)	0.002 (0.025)
常數項	1.429*** (0.047)	1.478*** (0.049)	1.495*** (0.050)	1.513*** (0.051)
var（受訪者層級常數項）	0.067*** (0.002)	0.073*** (0.002)	0.075*** (0.002)	0.078*** (0.002)
var（誤差項）	0.260*** (0.001)	0.286*** (0.001)	0.301*** (0.001)	0.332*** (0.001)
Log-likelihood	-153,608	-162,995	-168,085	-177,855
受訪者人數	3,875	3,875	3,875	3,875
觀察值個數	199,144	199,144	199,144	199,144

註：1. 括號內為標準誤。

2. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ 。

針對個人、題項層級控制變項，以下對於具顯著影響的變項加以說明。在題項層級變項方面，相對於事實題，態度題的題項回應時間較長，而其他主觀題的回應時間則較短，與敘述統計的結果（表 1）一致。至於個人層級變項，在教育程度方面，相對於高中職或以下教育程度者，專科、大學、研究所受訪者的題項回應時間均顯著較低。顯示教育程度愈高的受訪者，在解讀題項、萃取答案等方面的能力可能較低教育程度者為佳，而能以較快速的方式作答。然而，教育程度亦可能反映了時間的機會成本，教育程度較高的受訪者可能因為有工

作或事務較多，而在答題時以較快速的節奏填答。在會員招募管道方面，相較由手機簡訊網調加入者，由面訪、市話電訪、手機電訪成為會員的受訪者，題項回應時間顯著較高，顯示由手機簡訊網調加入者可能是時間成本較高的群體。

另外，表 3 顯示，在各控制變項對應變項的影響效果方面，無論是經 95%、98%、99% 臨界值處理後的資料，結果都與未經處理的分析結果相近，顯示分析結果對偏高數值的設定有相當的強韌性。對高於 95%、98%、99% 的題項回應時間，除了以 95%、98%、99% 臨界值取代外，本文也嘗試以遺漏值處理，或以題項回應時間的平均值、中位數取代。這三種資料處理方式（遺漏值處理、平均值取代、中位數取代）的分析結果，彙整於表 4。自表 4 可得知，主要分析結果與表 3 是相同的，表示主要分析結果對於題項回應時間離群值的處理方式具有強韌性。各控制變項的影響效果，亦與表 3 相近，為節省篇幅，未列於表 4 中。¹⁴

五、結論

注意力檢測已廣泛用於不少領域的網路調查中，但相關的方法研究仍不多見。在相關方法研究中，有關注意力檢測對受訪者作答行為、退出問卷的影響效果，儘管是重要的研究議題，研究仍相當罕見。

本研究採用機率樣本進行網路調查實驗，依照注意力檢測題的題目、放置位置區分為四個實驗組，並以未放置注意力檢測題的情境作為對照組，隨機將樣本分為五群，分析注意力檢測題對受訪者退出問

14 有興趣的讀者，可向作者索取相關分析結果。

表 4 題項回應時間自然對數值的多層級混合效果一般化線性模型：偏高數值以遺漏值處理、平均值取代，或中位數取代

	題項回應時間偏高數值處理方式								
	以遺漏值處理			以平均值取代			以中位數取代		
	(1) >95% 者 視為偏高值	(2) >98% 者 視為偏高值	(3) >99% 者 視為偏高值	(1) >95% 者 視為偏高值	(2) >98% 者 視為偏高值	(3) >99% 者 視為偏高值	(1) >95% 者 視為偏高值	(2) >98% 者 視為偏高值	(3) >99% 者 視為偏高值
注意力檢測群組									
A1	0.009 (0.013)	0.008 (0.014)	0.008 (0.015)	0.006 (0.013)	0.006 (0.014)	0.006 (0.015)	0.008 (0.013)	0.007 (0.014)	0.008 (0.015)
A2	0.003 (0.013)	0.000 (0.014)	0.002 (0.015)	0.004 (0.013)	0.002 (0.014)	0.004 (0.015)	0.004 (0.013)	0.001 (0.014)	0.004 (0.015)
B1	-0.018 (0.013)	-0.020 (0.014)	-0.020 (0.014)	-0.018 (0.012)	-0.020 (0.013)	-0.020 (0.014)	-0.018 (0.012)	-0.021 (0.013)	-0.020 (0.014)
B2	0.001 (0.013)	0.002 (0.014)	0.004 (0.014)	-0.001 (0.012)	0.001 (0.013)	0.003 (0.014)	-0.002 (0.012)	0.000 (0.013)	0.003 (0.014)
題項位置									
I(Item>17)	0.249*** (0.004)	0.249*** (0.005)	0.249*** (0.005)	0.252*** (0.004)	0.252*** (0.005)	0.253*** (0.005)	0.249*** (0.004)	0.250*** (0.005)	0.251*** (0.005)
I(Item>53)	0.266*** (0.006)	0.263*** (0.006)	0.262*** (0.006)	0.266*** (0.006)	0.263*** (0.006)	0.262*** (0.006)	0.271*** (0.006)	0.264*** (0.006)	0.261*** (0.006)
注意力檢測群組×題項位置									
A1·I(Item>17)	0.001 (0.006)	0.005 (0.006)	0.007 (0.007)	0.001 (0.006)	0.002 (0.006)	0.003 (0.007)	-0.001 (0.006)	0.001 (0.006)	0.001 (0.007)
B1·I(Item>17)	0.001 (0.006)	0.002 (0.007)	-0.000 (0.007)	0.000 (0.006)	0.002 (0.006)	-0.000 (0.007)	-0.001 (0.006)	0.002 (0.006)	-0.001 (0.007)
A2·I(Item>53)	0.087*** (0.010)	0.089*** (0.010)	0.089*** (0.010)	0.081*** (0.009)	0.087*** (0.010)	0.089*** (0.010)	0.078*** (0.009)	0.087*** (0.010)	0.089*** (0.010)
B2·I(Item>53)	0.070***	0.071***	0.069***	0.065***	0.069***	0.067***	0.064***	0.068***	0.067***
其他控制變項									
	V	V	V	V	V	V	V	V	V
Log-likelihood	-135,968	-149,709	-157,685	-142,830	-153,690	-160,448	-143,236	-153,960	-160,531
受訪者人數	3,874	3,875	3,875	3,874	3,875	3,875	3,874	3,875	3,875
觀察值個數	189,864	195,336	197,219	201,476	201,476	201,476	201,476	201,476	201,476

註：括號內為標準誤。*** $p<0.01$, ** $p<0.05$, * $p<0.1$ 。

卷與否、後續題項回應時間的影響效果。Cox 比例風險模型、logit 模型對退出問卷機率模型的分析結果均顯示，置有注意力檢測題的四個實驗組，在退出問卷風險（或部分填答機率）方面，與未放置注意力檢測題的對照組並無顯著差異，顯示設有注意力檢測題的問卷，不因注意力檢測題的題目為 *A* 或 *B*，以及在問卷中放置的位置，而影響到受訪者退出問卷的可能性。對於注意力檢測題是否會造成受訪者的反感，而對問卷完填帶來不利影響，在本研究的實驗設計之下，並未發現這樣的負面影響結果。

至於注意力檢測對題項填答時間的影響，根據既有文獻，可能的影響效果有正有負。一方面，注意力檢測題的放置，可能讓受訪者比較專注答題，而放緩後續題項的答題速度。另一方面，若設有注意力檢測題，受訪者可能因為注意力檢測題而覺得不受信任，因而加速後續題項的答題速度。¹⁵ 本研究採用兩層次模型的分析結果顯示，無論離群值是否經過處理以及如何選擇資料處理的臨界值，放置在後半部的注意力檢測題，會讓後續題項的回應時間顯著增加，表示注意力檢測題可能有讓受訪者花更多時間解讀題意、思考答案的效果。但若將注意力檢測題放置於問卷前半部，則在題項回應時間上，與未設注意力檢測題的對照組並無顯著差異。

前述注意力檢測題對退出問卷機率的影響效果分析，儘管研究設計與 Berinsky et al. (2016) 不同，仍有可相互參照之處。Berinsky et al. (2016) 的分析結果顯示，相較只進行一次性注意力檢測的對照組，重複施測的實驗組樣本退出問卷的可能性較高。不同於 Berinsky et al. (2016)，本文探討納入注意力檢測題與否對退出問卷機率的影響，並

15 參見 Oppenheimer et al. (2009)。

發現納入注意力檢測題的實驗組與未放置注意力檢測題的對照組，在退出問卷機率方面並無顯著差異。雖然本文的分析結果顯示，在網路調查問卷中納入注意力檢測題並不影響受訪者退出問卷的可能性，但 Berinsky et al. (2016) 警醒我們，如果在問卷中納入多道注意力檢測題，有可能對問卷完成帶來負面影響，顯示在注意力檢測題的使用上仍應慎重為之。

在注意力檢測題對後續作答行為影響效果的相關研究中，Hauser and Schwarz (2015) 發現注意力檢測題會讓受訪者在後續答題上更為審慎，Kung et al. (2018)、Gummer et al. (2021) 則發現注意力檢測題對後續答題行為並無顯著影響。本文的分析結果則顯示，注意力檢測題對題項回應時間是否具有顯著作用，與注意力檢測題的位置有關。在注意力檢測題放置於問卷後半部的情況下，注意力檢測題的設置，有增加後續題項回應時間的影響效果。這項分析結果在調查實務上的意涵是，如果研究者想在網路調查中藉由注意力檢測題提升受訪者的專注力，將注意力檢測題放在問卷後半部較前半部來得有效。

本研究不僅具有調查方法學理研究上的重要性，亦具有調查實務的意涵。有助於評估網路調查是否應設注意力檢測題，及納入注意力檢測題對受訪者作答行為、問卷完訪可能帶來的影響。儘管本研究具有這些貢獻，仍有些研究限制或未研析的課題有待進一步研究。首先，本文所設定的注意力檢測題是諸多題型中的一種，對於不同題型、不同問法、不同題目內容或不同選項設計的注意力檢測題，是否會影響受訪者退出問卷的可能性、後續題項回應時間，仍有待研究。第二，本文使用的樣本來自「網路調查會員資料庫」，加入並持續參與的會員可能與未加入或未持續參與者不同，在分析上是否存在樣本選擇偏誤 (sample selection bias)，有待進一步探究。第三，對於注意力檢測題

的影響效果，本文以問卷完填與否、後續題項回應時間作為應變項，但未分析注意力檢測題是否會影響受訪者繼續參與會員網路調查的意願，未來可進一步探討。第四，對於題項回應時間的離群值，本文已考慮多種資料處理方式，並發現在不同處理方式之下所得到的主要分析結果具有強韌性，但由於未能窮盡所有可能的處理方式，而現有文獻亦無指引可供依循，其他可能的離群值資料處理方式是否會帶來不同的結果，仍有待探究。最後，由於本文的研究目的在於探討注意力檢測對受訪者退出問卷與否、後續作答行為的影響，對受訪者通過注意力檢測與否，以及通過注意力檢測的受訪者是否真的比較專注，並未進行討論。然而，在分析資料的過程中，作者發現，在注意力檢測題置於問卷前半部的情況下，未通過者對注意力檢測題的回應時間，無論自平均值或中位數來看，均較通過者來得高，表示未通過者花在注意力檢測題上的時間較通過者為多。對注意力檢測題本身的回應行為、回應時間，都是值得未來深入探索的課題。

參考文獻

- Alvarez, R. M., L. R. Atkeson, I. Levin, and Y. Li, 2019, "Paying Attention to Inattentive Survey Respondents." *Political Analysis* 27(2): 145-162.
- Anduiza, E., and C. Galais, 2016, "Answering Without Reading: IMCs and Strong Satisficing in Online Surveys." *International Journal of Public Opinion Research* 29(3): 497-519.
- Atkeson, L. R., A. N. Adams, and R. M. Alvarez, 2018, "Mixing Survey Modes and Its Implications." Pp. 53-75 in *The Oxford Handbook of Polling and Survey Methods*, edited by L. R. Atkeson and R. M. Alvarez. New York: Oxford University Press.
- Bassili, J. N., and B. S. Scott, 1996, "Response Latency as a Signal to Question Problems in Survey research." *Public Opinion Quarterly* 60(3): 390-399.
- Berinsky, A. J., M. F. Margolis, and M. W. Sances, 2014, "Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-administered Surveys."

- American Journal of Political Science* 58(3): 739–753.
- 2016, “Can We Turn Shirkers into Workers?” *Journal of Experimental Social Psychology* 66: 20–28.
- Biemer, P. P., and S. L. Stokes, 1989, “The Optimal Design of Quality Control Samples to Detect Interviewer Cheating.” *Journal of Official Statistics* 5(1): 23–39.
- Callegaro, M., Y. Yang, D. S. Bholá, D. A. Dillman, and T. Y. Chin, 2009, “Response Latency as an Indicator of Optimizing in Online Questionnaires.” *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique* 103(1): 5–25.
- Clifford, S., and J. Jerit, 2015, “Do Attempts to Improve Respondent Attention Increase Social Desirability Bias?” *Public Opinion Quarterly* 79(3): 790–802.
- Couper, M. P., and F. Kreuter, 2013, “Using Paradata to Explore Item Level Response Times in Surveys.” *Journal of the Royal Statistical Society Series A (Statistics in Society)* 176(1): 271–286.
- Cox, D. R., and D. Oakes, 1984, *Analysis of Survival Data*. London: Chapman & Hall/CRC.
- Curran, P. G., 2016, “Methods for the Detection of Carelessly Invalid Responses in Survey Data.” *Journal of Experimental Social Psychology* 66: 4–19.
- Gummer, T., J. Roßmann, and H. Silber, 2021, “Using Instructed Response Items as Attention Checks in Web Surveys: Properties and Implementation.” *Sociological Methods & Research* 50(1): 238–264.
- Hauser, D. J., and N. Schwarz, 2015, “It’s a Trap! Instructional Manipulation Checks Prompt Systematic Thinking on “Tricky” Tasks.” *Sage Open* 5(2): 1–6.
- Ingram, D. D., and J. C. Kleinman, 1989, “Empirical Comparisons of Proportional Hazards and Logistic Regression Models.” *Statistics in Medicine* 8(5): 525–538.
- Johnson, M., 2004, “Timepieces: Components of Survey Question Response Latencies.” *Political Psychology* 25(5): 679–702.
- Jones, M. S., L. A. House, and Z. Gao, 2015, “Attribute Non-attendance and Satisficing Behavior in Online Choice Experiments.” *Proceedings in Food System Dynamics*: 415–432.
- Kreuter, F., S. Presser, and R. Tourangeau, 2008, “Social Desirability Bias in CATI, IVT, and Web Surveys: The Effects of Mode and Question Sensitivity.” *Public Opinion Quarterly* 72(5): 847–865.
- Krosnick, J. A., 1991, “Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys.” *Applied Cognitive Psychology* 5(3): 213–236.
- Kung, F. Y., N. Kwok, and D. J. Brown, 2018, “Are Attention Check Questions a Threat to Scale Validity?” *Applied Psychology* 67(2): 264–283.
- Leiner, D. J., 2019, “Too Fast, Too Straight, Too Weird: Non-reactive Indicators for Meaning-

- less Data in Internet Surveys.” *Survey Research Methods* 13(3): 229-248.
- Liu, M., & L. Wronski, 2018, “Trap Questions in Online Surveys: Results from Three Web Survey Experiments.” *International Journal of Market Research* 60(1): 32-49.
- Meade, A. W., and S. B. Craig, 2012, “Identifying Careless Responses in Survey Data.” *Psychological Methods* 17(3): 437-455.
- Olson, K., and J. D. Smyth, 2015, “The Effect of CATI Questions, Respondents and Interviewers on Response Time.” *Journal of Survey Statistics and Methodology* 3(3): 361-396.
- Oppenheimer, D. M., T. Meyvis, and N. Davidenko, 2009, “Instructional Manipulation Checks: Detecting Satisficing to Increase Statistical Power.” *Journal of Experimental Social Psychology* 45(4): 867-872.
- Read, B., L. Wolters, and A. J. Berinsky, 2021, “Racing the Clock: Using Response Time as a Proxy for Attentiveness on Self-administered Surveys.” *Political Analysis*. DOI: <https://doi.org/10.1017/pan.2021.32>.
- Silber, H., D. Danner, and B. Rammstedt, 2019, “The Impact of Respondent Attentiveness on Reliability and Validity.” *International Journal of Social Research Methodology*, 22(2): 153-164.
- Simon, H. A., 1956, “Rational Choice and the Structure of the Environment.” *Psychological Review* 63(2): 129-138.
- Tourangeau, R., L. J. Rips, and K. Rasinski, 2000, *The Psychology of Survey Response*. Cambridge, England: Cambridge University Press.
- Turner, G., P. Sturgis, and D. Martin, 2015, “Can Response Latencies Be Used to Detect Survey Satisficing on Cognitively Demanding Questions?” *Journal of Survey Statistics and Methodology* 3(1): 89-108.
- Yeager, D. S., J. A. Krosnick, L. Chang, H. S. Javitz, M.S. Levendusky, A. Simpser, and R. Wang, 2011, “Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-probability Samples.” *Public Opinion Quarterly* 75(4): 709-747.