The impact of careless responding on employability research

Inés Tomás [ID] **, Ana Hernández** [ID] **, Clara Cuevas** [ID] **, Vicente González-Romá** [ID]
IDOCAL, University of Valencia, Spain.

Abstract

Careless Responding (CR) compromises data quality in psychological and educational contexts. In this study we examine the impact of CR on research involving a key construct in Higher Education: personal employability. Particularly we assess how this impact depends on the strategy used to address CR by comparing four strategies: 1) using the total sample without taking any action, 2) eliminating careless respondents, 3) introducing CR as a control variable, 4) introducing CR as a moderating variable. Using a sample of 360 university graduates, results show that removing careless respondents reduces statistical power and some hypothesized employability effects become non-significant. In contrast, incorporating CR as a moderating variable was highly informative and yielded the best goodness-of-fit. These findings highlight the importance of considering CR behaviours and suggest that introducing CR as a moderator is an effective approach to maintaining data quality, statistical power and representativeness in employability research.

Keywords: careless responding; data quality; employability; young graduates.

1. Introduction

Personal employability is a critical construct in Higher Education, enhancing students' chances of securing quality jobs (González-Romá et al., 2018) and improving satisfaction (Vanhercke et al., 2016). Typically measured via questionnaires, employability research is vulnerable to Careless Responding (CR), a major source of bias that compromises data quality and validity. This study examines the indirect effects of employability on life satisfaction through core self-evaluations while investigating how different strategies for managing CR impact tested relationships and model fit. Four approaches are compared: 1) using the full sample without addressing CR, 2) excluding careless respondents, 3) incorporating CR as a control variable,

and 4) treating CR as a moderating variable. Findings will shed light on mechanisms linking employability and satisfaction while offering practical recommendations to address CR and safeguard validity in employability research.

1.1. Addressing and managing Careless Responding

Careless responding (CR) occurs when participants fail to thoroughly read or focus on items' content, producing unreliable data that misrepresents their true construct levels (Ward & Meade, 2022). Among various CR detection methods (see Edwards, 2019, Ward & Mead, 2022), this study focuses on Instructed Response Items (IRIs), a widely recommended approach (e.g., Berinsky et al., 2021; Maniaci & Rogge, 2014; Meade & Craig, 2012). IRIs instruct participants to select a specific response (e.g., "5. Completely agree"), with failure to comply indicating careless responding. This approach stands out for its simplicity, transparency, and strong metric properties (Kam & Chan, 2018).

Despite concerns about CR's negative impact on research validity (Kim & Oh, 2022; Maniaci & Rogge, 2014), there is no consensus on how best to address it. While excluding careless respondents is common (e.g., Maniaci & Rogge, 2014; Meade & Craig, 2012), this reduces sample size, lowers statistical power, and risks losing representativeness by excluding unique participants (e.g., Goldammer et al., 2020; Ward & Meade, 2022). Alternatives include treating CR as a control (Goldammer et al., 2020) or moderating variable (Edwards, 2019). However, no published research has compared the effects of these strategies. This study fills the gap by examining the most effective methods for managing CR, focusing on employability research.

1.2. From employability to well-being: A mediated model through core-self evaluations

In this study we focus on personal employability. It is defined as "a form of work-specific active adaptability that enables workers to identify and realize career opportunities" (Fugate et al. (2004; p. 16). It is a multidimensional construct that encompasses four personal factors: a) career identity, which reflects one's sense of self and aspirations within a career, serving as a motivational driver; b) personal adaptability, or the ability to respond to environmental demands and explore opportunities; c) human capital, including attributes such as education, work experience, and skills; and d) social capital, the advantages gained through social networks that provide access to information and opportunities.

In the employability literature, a central hypothesis is that employability enhances satisfaction and well-being. It fosters a sense of control over one's employment situation, positively influencing outcomes like job satisfaction and life satisfaction (De Cuyper et al., 2011). Additionally, job insecurity has been identified as a mediating mechanism linking employability and well-being indicators such as engagement and life satisfaction (De Cuyper et al., 2008). In

this study, we propose core self-evaluations (individual's subconscious, fundamental evaluations about themselves) as the mediator between employability and life satisfaction.

1.3. Hypotheses

Our research model is depicted on the left panel of Figure 1. All relations are expected to be positive. Regarding the best way of managing CR, we hypothesize that using the full sample without addressing CR will be the least effective option. The limited research on this topic prevents us from making a specific hypothesis about which of the other strategies (excluding careless respondents, including CR as a control variable, introducing CR as a moderator), is most effective.

2. Method

2.1. Participants and procedure

This study was part of a larger investigation on graduates' overqualification in Spain. To test our hypotheses, we utilized a sample of undergraduate and master's students that were about to complete their studies. Participants were first contacted approximately one month before graduation (T1), and were informed they would be invited to complete two additional questionnaires: nine months post-graduation (T2) and four months after T2 (T3). The final sample consisted of 360 participants who completed the online questionnaire at all three time points. In terms of educational level, 61.1% had earned a bachelor's degree, while 38.9% had completed a master's degree. Most participants came from Social Sciences (53.6%) and Health (23.6%). The average age was 25.6 years (SD = 6.3), and 71.7% were female. Careless respondents were those who made one or more errors on the three IRIs included in the questionnaire. At T1, 8.3% of participants were identified as careless respondents, increasing to 14.4% at T2, and 14.7% at T3.

2.2. Measures

Careless responding was measured by means of three IRIs that were embedded within different scales of the questionnaire, each specifying the response to be selected. For instance, in a scale ranging from 1 ("Never") to 6 ("Always"), an IRI example would read: "To control the quality of survey responses, now select the option 'never'". We computed the total number of errors (responses different from the ones specified in the IRIs) across the three IRIs (from 0 to 3).

Employability was measured at T1 including the four factors proposed by Fugate et al. *Career identity* was measured with González-Romá et al.'s (2018) 4-item scale (e.g., "I identify with the line or area of work that I have chosen"), rated on a 6-point Likert scale (1. Strongly Disagree, 6. Strongly Agree). *Personal Adaptability* was measured with a 3-item scale

developed for this study (e.g., "I am able to adapt to the changing circumstances of my environment"). Items were rated on a 5-point scale (1. Not at all, 5. A lot). *Human Capital* was measured with a 6-item scale based on Hernández-March et al.'s (2009) generic competences: oral and written communication, problem solving, time and resource management, teamwork, continuous learning, and taking responsibility. Items were rated on a 5-point Likert scale (1. Low degree, 5. High degree). *Social Capital* was measured by means of a 4-item scale based on González-Romá et al. (2018) (e.g., "I have a network of professional contacts that will help me find job opportunities"). Items were rated on a 6-point Likert scale (1. Strongly Disagree, 6. Strongly Agree). Confirmatory factor analysis supported the expected four-factor solution (χ^2 =250, df = 113, p < .001; RMSEA = 0.06; SRMR = 0.05; CFI = 0.94; TLI = .93) with factor loadings ranging between .41 and .91 (p<.001). Cronbach's alphas ranged between .72 and .85.

Core-self evaluations (CSE) were measured at T2. To measure CSE, we used the Core Self-Evaluations Scale developed by Judge et al. (2003). The scale consists of 12 items (e.g., "I am confident I will get the success I deserve in life") that are responded to on a 6-point Likert scale (1. Strongly Disagree, 6. Strongly Agree). Cronbach's alpha value was .85.

Life satisfaction was assessed at T3 by the five-item Satisfaction with Life Scale (Diener et al., 1985) (e.g., "In most ways my life is close to my ideal."). Cronbach's alpha was .88.

2.3. Analysis

We tested the proposed mediation model by means of Structural equation modeling (SEM). We first tested the model on both the total sample (M1) and the clean sample (M2) (after excluding careless respondents). Additionally, two models were tested using the full sample while accounting for "careless responding" in two ways: as a control variable influencing the relevant substantive variables (M3), and as a moderator of all proposed substantive relationships (M4). The goodness-of-fit of the four models was compared. Differences of no more than .015 in RMSEA values (ΔRMSEA; Chen, 2007) and no more than 0.01 in CFI and TLI values (ΔCFI and ΔTLI; Cheung & Rensvold, 2002; Widaman, 1985) were considered indicative of negligible practical differences. Analyses were conducted using SPSS 28 and Mplus version 8.8.

3. Results

3.1. Goodness-of-fit of the models

Table 1 shows the goodness-of-fit comparison across the alternative models tested. The best fitting model was the one that introduced CR as a moderator (M4). M4 showed relevant differences compared to models M1 (ΔRMSEA = .047, ΔCFI = .025, ΔTLI = .057), M2 (ΔRMSEA = .051, ΔCFI = .027, ΔTLI = .062), and M3 (ΔRMSEA = .036, ΔCFI = .038, ΔTLI = .077). Conversely, the model with CR as a control variable (M3) exhibited the poorest fit,

showing relevant differences with both M1 (ΔCFI = .013, ΔTLI = .020), and M2 (ΔCFI = .011, ΔTLI = .015). The differences between M1 (full sample) and M2 (sample without careless respondents) were negligible in terms of fit (ΔRMSEA = .004, ΔCFI = .002, ΔTLI = .005).

Table 1. Goodness-of-fit of the alternative models tested.

Model	χ^2	df	RMSEA	CFI	TLI	SRMS
M1: Full sample (N=360)	11.01*	4	0.070	0.967	0.925	0.040
M2: Eliminating CR (N=248)	9.42	4	0.074	0.965	0.920	0.041
M3: CR as control variable (N=360)	36.30**	16	0.059	0.954	0.905	0.041
M4: CR as moderating variable (N=360)	13.03	11	0.023	0.992	0.982	0.029

Note: CR = careless responding. *p<.05; **p<.01.

3.2. Relationships between the variables

Figures 1 and 2 illustrated that the regression coefficients across the tested models were largely consistent. All hypothesized relationships were statistically significant, except for M2 (the model excluding careless respondents), where one of the tested relationships and the corresponding indirect effect were not statistically significant. Specifically, while all four employability factors showed positive and statistically significant relationships with CSE, the relationship between social capital and CSE in M2 was not statistically significant (β=.13, p=.055). Similarly, indirect effects of the employability factors on life satisfaction through CSE were significant (as the 95% bias-corrected (BC) bootstrap confidence interval (CI) excluded zero), with values ranging from .08 to .25, except for the indirect effect of social capital on life satisfaction in M2, which was not significant (IE = .07; BC bootstrap 95% CI = [-.001, .15]).
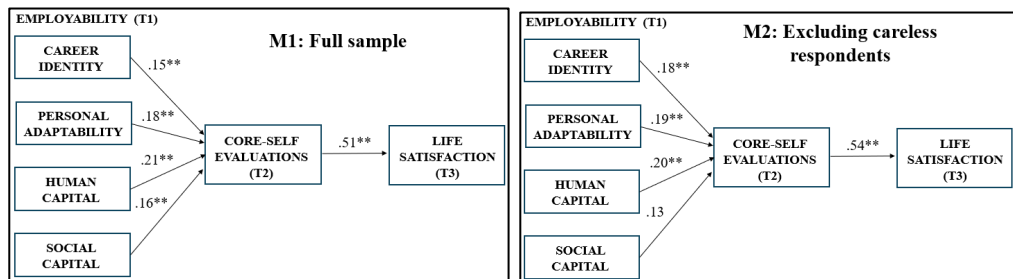
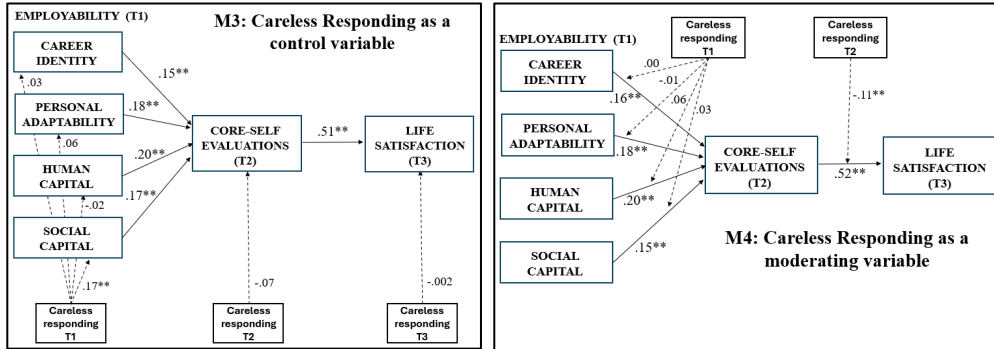*Figure 1. Standardized regression coefficients for M1 and M2.*

Figure 2. Standardized regression coefficients for M3 and M4.

Interestingly, in M4 one of the moderating effects of CR was statistically significant. This significant interaction term indicates that the effect of CSE on life satisfaction differs across groups of respondents defined by the number of errors in the IRIs. To further examine the interaction effect, we used Process macro for SPSS (Hayes, 2013) to compute simple slopes for the four levels of the moderator: careful (0 errors) and careless respondents (1, 2, or 3 errors). We also plotted the corresponding regression lines (see Figure 3). The results showed that the slope estimating the relationship between CSE and life satisfaction was positive and statistically significant for careful respondents (b=0.84, p<.001) and for careless respondents with 1 error (b=0.51, p<.001). However, the relationship was not statistically significant for careless participants with 2 errors (b=0.17, p=.51) and 3 errors (b=-0.16, p=.69).
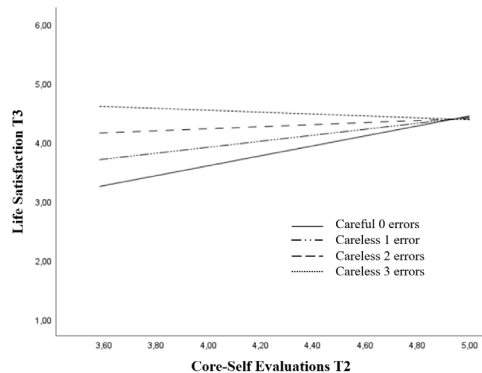

Figure 3. Relationship between CSE and life satisfaction as a function of careless responding.

4. Discussion

The results reveal that all factors in Fugate et al.'s personal employability model contribute to enhancing well-being (life satisfaction) through the increase of CSE. However, this indirect

effect was not supported when careless respondents were excluded from the sample, as the relationship between social capital and CSE became statistically nonsignificant. These findings highlight the impact of different strategies for managing CR on employability research outcomes. Specifically, the results suggest that in studies with relatively small sample sizes (360 in this study), excluding careless respondent reduces statistical power, potentially obscuring significant relationships. Additionally, using CR as a control variable appears to be the less advisable option, as it yielded the poorest model fit. On the other hand, incorporating CR as a moderating variable proved to be the most effective strategy, offering the best model fit while preserving statistical power by retaining the full sample. Analysis of the moderating effect of CR showed that the relationship between CSE and life satisfaction was stronger for careful respondents, becoming nonsignificant for groups with higher errors in the IRIs.

Several limitations of our study should be acknowledged. Contextual factors such as sample size, the proportion of careless respondent, and the length of the questionnaire remained constant, which restricts the generalizability of our findings to different conditions. Future simulation studies are needed to explore how these and other contextual variables may influence the impact of CR on psychological research outcomes.

Despite these limitations, our findings have valuable theoretical and practical implications. By comparing various CR management strategies, we recommend against excluding careless respondents. Instead, we emphasize the importance of accounting for CR behaviors during data analysis. Introducing CR as a moderating variable emerges as an effective strategy for preserving data quality and representativeness in employability research.

References

Berinsky, A. J., Margolis, M. F., Sances, M. W., & Warshaw, C. (2021). Using screeners to measure respondent attention on self-administered surveys: Which items and how many? *Political Science Research and Methods*, 9(2), 430-437. doi:10.1017/psrm.2019.53

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14, 464-504. https://doi.org/10.1080/10705510701301834.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233-255. https://doi.org/10.1207/s15328007sem0902_5.

De Cuyper, N., Bernhard-Oettel, C., Berntson, E., De Witte, H., & Alarco, B. (2008). Employability and employees' well-being: Mediation by job insecurity. Applied Psychology: An International Review, 57(3), 488–509. https://doi.org/10.1111/j.1464-0597.2008.00332.x.

De Cuyper, N., Van der Heijden, B. I. J. M., & De Witte, H. (2011). Associations between perceived employability, employee well-being, and its contribution to organizational success: a matter of psychological contracts? The International Journal of Human Resource Management, 22(7), 1486–1503. https://doi.org/10.1080/09585192.2011.561962.

Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49, 71–75. https://doi.org/10.1207/s15327752jpa4901_13.

Edwards, J. R. (2019). Response invalidity in empirical research: Causes, detection, and remedies. *Journal of Operations Management*, 65(1), 62-76. https://doi.org/10.1016/j.jom.2018.12.002.

Fugate, M., Kinicki, A. J., & Ashforth, B. E. (2004). Employability: A psycho-social construct, its dimensions, and applications. *Journal of Vocational Behavior*, 65(1), 14–38. https://doi.org/10.1016/j.jvb.2003.10.005.

Goldammer, P., Annen, H., Stöckli, P. L., & Jonas, K. (2020). Careless responding in questionnaire measures: Detection, impact, and remedies. The Leadership Quarterly, 31(4), Article 101384. https://doi.org/10.1016/j.leaqua.2020.101384.

González-Romá, V., Gamboa, J. P., & Peiro, J. M. (2018). University graduates' employability, employment status, and job quality. *Journal of Career Development*, 45(2), 132-149. https://doi.org/10.1177/0894845316671607.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press. https://doi.org/10.1111/jedm.12050.

Hernández-March, J., Martín del Peso, M., & Leguey, S. (2009). Graduates' skills and higher education: The employers' perspective. *Tertiary Education and Management*, 15(1), 1-16. https://doi.org/10.1080/13583880802699978.

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The Core Self-Evaluations Scale: Development of a measure. Personnel Psychology, 56(2), 303–331. https://doi.org/10.1111/j.1744-6570.2003.tb00152.x.

Kam, C. C. S., & Chan, G. H.-h. (2018). Examination of the validity of instructed response items in identifying careless respondents. Personality and Individual Differences, 129, 83–87. https://doi.org/10.1016/j.paid.2018.03.022.

Kim, N., & Oh, J. (2022). Learning from mistakes: Impact of careless responses on counseling research using amazon's mechanical turk. Measurement and Evaluation in Counseling and Development. Advance online publication. doi: 10.1080/07481756.2022.2073895.

Maniaci, M. R., & Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. Journal of Research in Personality, 48, 61–83. https://doi.org/10.1016/j.jrp.2013.09.008.

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. Psychological Methods, 17(3), 437–455. https://doi.org/10.1037/a0028085.

Vanhercke, D., De Cuyper, N., & De Witte, H. (2016). Perceived employability and well-being: An overview. *Psihologia Resurselor Umane*, 14(1), 8-18.

Ward, M. K., & Meade, A. W. (2022). Dealing with careless responding in survey data: Prevention, identification, and recommended best practices. *Annual Review of Psychology*, 74(1), 577-596. https://doi.org/10.1146/annurev-psych-040422-045007.

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement*, 9(1), 1-26. https://doi.org/10.1177/014662168500900101.