Quantitative Data Analysis Critical Review:

Paper Citation: Quintana, D., Cervantes, A., Saez, Y., & Isasi, P. (2018). Internet Use and Psychological Well-Being at Advanced Age: Evidence from the English Longitudinal Study of Aging. International Journal of Environmental Research and Public Health, 15(3).
Journal Subject Area: Environmental Science and Medicine. H-Index: 201.

Details of Approach:

Quintana *et al*. (2018) fundamentally address the question of to what extent internet usage, particularly email communications, is associated with different dimensions of psychological well-being in adults over the age of 50. The paper builds on various prior studies which utilised similar statistical analysis techniques and/or data sets, including Nie *et al.* (2017), Xavier *et al.* (2014), Jivraj *et al*., (2014) *inter alia*. Instead of formulating a direct *research question*, the *thesis statement* outlines that "[t]his work explores the connection between psychological well-being and Internet use in older adults" (1), with the study later using regression coefficients, p-values, and confidence intervals to communicate the relationship between internet usage and psychological well-being in elder populations after accounting for various covariates.

The study uses the English Longitudinal Study of Aging (ELSA), a dataset with "information on the dynamics of health, social, well-being and economic circumstances in the English population aged 50 and older" (ELSA, n.d.: n.p.). The study began in 2002, with follow-up studies occurring biennially *[every two years]*. At recruitment, the sample consisted of 12,099 participants, with 11,391 core members and 708 subsidiary partners; 72 of which were aged ≥50 years (Steptoe, 2017). Quintana *et al*.'s study specifically uses waves three through seven, spanning from 2006 to 2015.

The datasets were primarily produced through in-person, face-to-face interviews, with only ≈2% of interviews being undertaken by proxy interviewees *(ibid.)*. Computer-assisted interviewing techniques were adopted for speed and functionality (Steptoe *et al.*, 2013), with additional questionnaires given on pen and paper. This is supplemented by the collection of bio-measures near to the interview time. ELSA measures a number of core variables through seven key domains. There are various examples of nominal, ordinal and interval-/ratio level variables within the dataset, such as: sex *[nominal, dichotomous]*, self-rated health status *[ordinal, measured 0-4]*, household assets *[interval-level variable, equal intervals and no true '0-point'] inter alia*. Additional derived variables are also formed from the core domains, with the derived variable 'CASP-19' frequently used in Quintana *et al*.'s study as a metric of 'Quality of Life'.

When analysing the data, Quintana *et al*. outline three key domains which constitute psychological well-being. These are derived variables which have been discussed and legitimised in prior literature (*cf.*: Steptoe *et al.*, 2012; Jivraj *et al*., 2014; Lifshitz, 2018): 'evaluative component' *[satisfaction with life scale]*, 'hedonic component' *[CASP-19]*, and 'eudaimonic component' *[CASP-19]*. These indicators of psychological well-being are sustained throughout the data analysis and indicate different 'levels' of psychological security. In relation to internet usage, "[t]he main predictor was the reply to the [question] 'I use

Internet/E-mail: yes/no'. This proxy for digital literacy was represented with a dichotomous variable" *[quasi-nominal level variable]* (Quintana *et al*., 2018: 3). In this study, confounding variables were considered to ensure that the relationship between internet use and psychological well-being was accurately assessed. Confounding variables are variables that can affect both the independent variable *[internet use]* and the dependent variable *[psychological well-being]*, potentially distorting or masking the actual relationship between them (Martin. 2000). Mobilising such identification of various sociodemographic and health-related variables that could influence both internet use and well-being, the study uses regression models to include these variables as covariates, allowing for adjustments. By doing so, the later analysis is able to isolate the effect of internet use on well-being, filtering out the confounding influence of these variables to produce more accurate conclusions. Summary statistics are present first, before the Generalised Estimating Equations (GEE) statistical test is applied. Data is presented in tables throughout the paper.

Summary Statistics:

Table 1. Baseline characteristics of the analytical sample measured at Wave 3. English Longitudinal Study of Aging 2006–2007. Main descriptive statistics.

	Mean	Std. Dev.	Min.	Max.
SWLS Score *	20.49	6.159	0	30
EOLS Score *	10.13	1.624	2	12
EDS Score *	32.94	6.589	6	45
Internet/Email User	0.66	0.475	0	1
Delayed Recall	5.32	1.778	0	10
Physical Activity	2.09	0.719	0	3
Org. membership	1.79	1.410	0	8
Voluntary Work	0.39	0.487	0	1
Sex	0.55	0.498	0	1
Marital Status	0.75	0.431	0	1
Education	1.18	0.807	0	2
Lack of impairments	0.85	0.360	0	1
Wealth Quintile	3.35	1.357	1	5
Age Interval	1.74	0.780	1	4

* Scales used to measure the tree core components of psychological well being: evaluative (SWLS), hedonic (EOLS) and eudaimonic (EDS).

Table 2. Baseline characteristics of the analytical sample measured at Wave 3 by Internet/Email use. English Longitudinal Study of Aging 2006–2007. Main descriptive statistics.

	Non-Internet/Email Users				Internet/Email Users *			
	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.	Min.	Max.
SWLS Score *	19.77	6.368	0	30	20.86	6.015	0	30
EOLS Score *	9.93	1.781	2	12	10.24	1.526	4	12
EDS Score *	31.20	6.983	8	45	33.59	6.280	6	45
Delayed Recall	4.79	1.748	0	10	5.60	2.989	0	10
Physical Activity	1.97	0.732	0	3	2.16	0.703	0	3
Org. membership	1.52	1.339	0	8	1.93	1.426	0	8
Voluntary Work	0.34	0.224	0	1	0.41	0.492	0	1
Sex	0.62	0.485	0	1	0.51	0.5	0	1
Marital Status	0.68	0.467	0	1	0.79	0.165	0	1
Education	0.78	0.028	0	2	1.40	0.732	0	2
Lack of impairments	0.77	0.419	0	1	0.89	0.319	0	1
Wealth Quintile	2.94	1.345	1	5	3.56	1.315	1	5
Age Interval	2.06	0.821	1	4	1.57	0.701	1	4

* Scales used to measure the tree core components of psychological well being: evaluative (SWLS), hedonic (EOLS) and eudaimonic (EDS).

The summary statistics act as the baseline for later GEE tests, additionally highlighting the various baseline characteristics of the sample. Quintana *et al*. chose to apply the GEE test, which is a method for analysing longitudinal and correlated data, particularly where repeated measurements are taken on the same subjects over time; such as in the ELSA.

Correlated data refers to data points that are not independent of each other, but rather show some degree of relationship or association. This means that the observations are not fully random or independent, which affects how the data should be analysed. There are 3 reasons why this data might be correlated. Firstly, repeated measures, the same subjects are measured multiple times over a period. Each individual's measurements tend to be similar across time points because they are influenced by that person's unique characteristics. Secondly, clustered data, observations can also be clustered within groups like families, regions or age *[over 50]*. Measurements within the same group tend to be more similar than those between different groups due to shared environments, genetic factors or socio-cultural influences. Finally, cross-correlations, different variables can also be correlated if they are affected by common factors. In longitudinal studies, such as Quintana *et al*.'s paper, repeated measures of the same individuals result in correlated data. GEEs account for this inter-subject correlation, ensuring that statistical inferences are accurate.

By applying the GEE as a significance test, different regression coefficients, p-values, and confidence intervals are obtained from the data set that account for the various covariates.

To specifically communicate their results, *p values* are discussed to show the strength of evidence against the null hypothesis *[there is no relationship between internet usage and psychological well-being, h_0]*, thus indicating the probability of obtaining the observed results. The study espouses that "[t]he impact of digital literacy, despite of being more modest than the one observed for physical activity, lack of functional impairments or wealth, is still relevant even after controlling for all covariates" (9), as indicated by the evaluative *p*-value *[0.078]* and the eudaimonic *p*-value *[0.015]*. The hedonic *p*-value was insignificant at 0.192. Moreover, the regression coefficients indicate a positively strong relationship between the evaluative and eudaimonic aspects of well-being, at 0.367 and 0.584 respectively. Again, the hedonic regression coefficient was weaker, at only 0.075. These values align with prior studies, with additional covariates supporting the findings of wider research into well-being and fitness, volunteering *inter alia*.

Results:

	SWLS Score *				EOLS Score *				EDS Score *			
	Coef. †	95% Conf. Int.		p §	Coef. †	95% Conf Int.		p §	Coef. †	95% Conf. Int.		p §
		Inf	Sup			Inf	Sup			Inf	Sup	
Internet/Email user (predictor)				0.078				0.192				0.015
No	Ref				Ref				Ref			
Yes	0.367	−0.041	0.776		0.075	−0.037	0.187		0.584	0.113	1.054	
Physical Activity				<0.001				<0.001				<0.001
Sedentary	Ref				Ref				Ref			
Low	0.496	−0.056	1.048		0.274	0.128	0.420		0.970	0.377	1.562	
Moderate	1.023	0.464	1.583		0.570	0.418	0.722		2.139	1.528	2.750	
High	1.448	0.812	2.084		0.804	0.634	0.975		3.153	2.458	3.848	
Voluntary Work				<0.001				0.001				0.001
No	Ref				Ref				Ref			
Yes	0.938	0.606	1.270		0.158	0.068	0.247		0.601	0.233	0.968	
Sex				0.633				<0.001				0.120
Male	Ref				Ref				Ref			
Female	0.098	−0.303	0.499		0.226	0.116	0.336		0.354	−0.092	0.801	
Marital Status				<0.001				<0.001				0.723
Single	Ref				Ref				Ref			
Married	2.250	1.741	2.760		0.253	0.125	0.382		−0.094	−0.618	0.429	
Education				0.001				0.136				0.746
None	Ref				Ref				Ref			
Intermediate	−1.016	−1.556	−0.477		0.066	−0.081	0.214		−0.170	−0.784	0.443	
Degree	−0.737	−1.281	−0.193		−0.061	−0.216	0.094		−0.241	−0.862	0.379	
Lack of impairments				<0.001				<0.001				<0.001
Yes	Ref				Ref				Ref			
No	2.613	2.155	3.071		0.897	0.777	1.016		4.535	4.041	5.029	
Wealth Quintile				<0.001				<0.001				<0.001
Q1 (Poorest)	Ref				Ref				Ref			
Q2	1.102	0.416	1.787		0.202	0.028	0.377		1.282	0.525	2.038	
Q3	1.353	0.672	2.035		0.334	0.160	0.509		1.794	1.030	2.557	
Q4	2.028	1.336	2.720		0.455	0.278	0.631		2.883	2.119	3.647	
Q5 (Wealthiest)	2.766	2.068	3.463		0.463	0.285	0.642		3.611	2.838	4.384	

Critique of Approach:

Through its robust and appropriate usage of GEEs to analyse the between internet use and psychological well-being, there are many advantages to the statistical method adopted by Quintana *et al*. Fundamentally, GEEs provide an estimate of the average effect of internet use on well-being across the population, which is particularly useful for public health and policy recommendations. Moreover, by breaking the complex notion of psychological well-being into three comparable, easily categorised variable allows for more detailed analysis and targeted policy implementation. This is especially clear in the vast differences between evaluative/eudaimonic and hedonic aspects of well-being in the elderly. Holistically, their appropriate use of GEEs on time-series data produced statistics that are in line with prior research findings whilst opening new avenues of investigation through their deconstructionist approach. However, that is not to say that there aren't disadvantages to the statistical method which was adopted.

Some of these disadvantages are explicitly discussed in the paper, such as the lack of information on 'frequency of use' and sampling bias towards non-care home residents. Nevertheless, there remain some more fundamental disadvantages to their approach which are enforced by the statistical method. For example, GEEs rely on assumed correlation structures between variables, meaning that the actual relationship between variables may not be accurately captured (Ballinger, 2004; Hanushek & Jackson, 2013). Even though the study accounted for covariate adjustments/influence, GEEs focus on population-level effects and don't directly capture individual-level variations. Consequently, the causal interpretation of the data is obscured, with correlational findings taking precedence. As an alternative approach, the study *could* have utilised matching to manage for the potential selection bias present in data which has numerous confounding variables.

Matching is a statistical method to address potential selection bias when dealing with data that has a lot of potential confounding variables/covariates and is grouped by different characteristics. By using Propensity Score Matching (PSM), comparable groups of internet users and non-users can be matched on comparable demographic, socio-economic and health characteristics. The effect of internet use on well-being could then be assessed by comparing these matched pairs (Davis et al., 2021; Fantechi & Fratesi, 2022).

Such an approach presents an alternative methodological direction to the chosen analytical framework, as GEEs solely focus on population-averaged effects. Conversely, PSM can evaluate treatment effects within more specific population subsets *[matched pairs]* to understand nuanced impacts and make treatment comparisons clearer and directly observable; thus allowing policymakers to understand how different groups may benefit differently from internet use. As the effect of internet use on well-being is more directly observable in these matched pairs/groups, the outcome difference is less likely to be influenced by external factors, reducing the selection bias associated with GEEs. Moreover, by matching individuals with similar propensity scores *[probability of internet use based on covariates]*, PSM mimics the balance achieved in randomised controlled trials. This helps isolate the effect of internet use on psychological well-being.

References:

Ballinger, G. A. (2004). Using generalized estimating equations for longitudinal data analysis. Organizational research methods, 7(2), 127-150.

Davis, M. L., Neelon, B., Nietert, P. J., Burgette, L. F., Hunt, K. J., Lawson, A. B., & Egede, L. E. (2021). Propensity score matching for multilevel spatial data: accounting for geographic confounding in health disparity studies. International Journal of Health Geographics, 20, 1-12.

ELSA. (n.d.). About ELSA. https://www.elsa-project.ac.uk/about-elsa

Fantechi, F., & Fratesi, U. (2022). Measuring competitiveness differentials inside the same region: a propensity-score matching approach. Social Indicators Research, 1-27.

Hanushek, E. A., & Jackson, J. E. (2013). Statistical methods for social scientists. Academic Press.

Jivraj, S., Nazroo, J., Vanhoutte, B., Chandola, T. (2014). Aging and Subjective Well-Being in Later Life. The Journals of Gerontology: Series B, 69(6), 930–941.

Lifshitz, R., Nimrod, G., Bachner, Y.G. (2018). Internet use and well-being in later life: A functional approach. *Aging & Mental Health, 22*, 85–91.

Martin, D. W. (2000). Doing psychology experiments. Wadsworth/Thomson Learning.

Nie, P., Sousa-Poza, A., & Nimrod, G. (2017). Internet Use and Subjective Well-Being in China. Social Indicators Research, 132, 489–516.

Quintana, D., Cervantes, A., Saez, Y., & Isasi, P. (2018). Internet Use and Psychological Well-Being at Advanced Age: Evidence from the English Longitudinal Study of Aging. International Journal of Environmental Research and Public Health, 15(3).

Steptoe, A. (2017). ELSA [Data set]. Dementias Platform UK.

Steptoe, A., Breeze, E., Banks, J., Nazroo, J. (2013). Cohort Profile: The English Longitudinal Study of Ageing. *International Journal of Epidemiology, 42*(6), 1640–1648.

Steptoe, A., Panayotes, P., de Oliveira, C. (2012). The psychological well-being, health and functioning of older people in England. In The Dynamics of Ageing: Evidence from the English Longitudinal Study of Ageing 2002–10 (Wave 5). Institute for Fiscal Studies: London, UK.

Xavier, A. J., d'Orsi, E., de Oliveira, C. M., Orrell, M., Demakakos, P., Biddulph, J. P., Marmot, M. G. (2014). English Longitudinal Study of Aging: Can Internet/E-mail Use Reduce Cognitive Decline? The Journals of Gerontology: Series A, 69(9), 1117–1121.