



Electronic Journal of Business Research Methods (EJBRM)

Volume 20 Issue 2 (2022)

Edited by Ann Brown

Contents Page

1. Professor Vincent **Cassar**, Professor Frank **Bezzina**, Dr Stephanie **Fabri**; *A Fuzzy-set Approach to Re-Exploring Work-related Stress-outcome Relationships: Implications for Research Methods, Theory and Practice*; pp63-74
2. Aliosha **Alexandrov**, Sandipan **Sen**, Michael **Tippins**; *Network Analysis of Traditional Word of Mouth*; pp75-88
3. Faies **Jafar**; *Applicability of Grounded Theory Methodology in International Relations Studies*; pp89-99

A Fuzzy-set Approach to Re-Exploring Work-related Stress-outcome Relationships: Implications for Research Methods, Theory and Practice

Vincent Cassar, Frank Bezzina and Stephanie Fabri

Dept. of Management, Faculty of Economics, Management and accountancy, University of Malta, MALTA

Vincent.cassar@um.edu.mt

Frank.bezzina@um.edu.mt

Stephanie.fabri@um.edu.mt

Abstract: This study employs fuzzy-set Qualitative Comparative Analysis to re-examine the more traditional correlational and regression models used to understand work stress and outcome relationships. Indeed, theoretical frameworks have been constructed underlying this premise and so have interventions assumed such principles. This study argues that the relationship between perception of stressors is not necessarily continuously linear and hence assessing non-linear but also more complex configurations becomes pivotal. While it is acknowledged that theory is important to simplify the explanation underlying the phenomenon of interest, likewise theory should also acknowledge new patterns that hide 'inside' the data and which reveal new pockets of information that explain the phenomenon better. Therefore, in line with calls for researchers to explore further new configurational patterns in the data and thus help build and refine theory, we examine potential complex associations between work-related stress and salient outcomes namely mental health, physical symptoms, presenteeism and intention to quit using fuzzy-set Qualitative Comparative Analysis instead of the more traditional correlational approach. Fuzzy-set methodology involves a number of parameters that include non-linearity but also equifinality and asymmetric associations. Responses were obtained from 509 employees who volunteered to fill in a work-related stress questionnaire as part of a wider campaign related to the management of psychosocial hazards at work and the data was subject to fuzzy-set configuration analysis between stressors and outcomes. Fuzzy-set analyses revealed that specific points of high or low levels of work stress did not necessarily concur with more common regression results but revealed interesting pattern associations between work related stress and outcomes. The results imply that by adopting traditional methods, the relations observed would be either over- or under- estimated and researchers are urged to explore new research methodological approaches such as fuzzy-set methodology in the analyses of their data that may reveal new patterns and hence better inform both theory and practice.

Keywords: fuzzy-set Qualitative Comparative Analysis; work-related stress; stress interventions; complexity; non-linearity; equifinality; asymmetry; evidence-based practice.

1. Introduction

The management of stress at work has been very high on the international agenda for over two decades or so (Cox, Griffith and Rial-Gonzalez, 2000; Leka, Cox and Zwetsloot, 2008; Hassard et al., 2018). A number of international bodies (e.g. the European Agency for Safety and Health at Work (EU-OHSA)) have provided very interesting insights on the most important emerging psychosocial hazards, and many opine that work stress can be better managed or prevented if properly measured. Several tools have been developed to this end and this endeavour is part of a wider framework to tackle psychosocial hazards, largely based on the risk management paradigm (Kortum, Leka and Cox, 2011). These tools offer the possibility for the development of a social dialogue that is critical in defining the best approach to work-stress evaluation, management and prevention (Zoni and Luchini, 2012). Psychosocial hazards have visible consequences on a number of salient outcomes not least employees' physical and psychological health (van Stolk, et al., 2014).

To improve practices that render workplaces less stressful and more conducive to enhanced well-being, researchers and practitioners alike need to understand the current positioning of the field in order to identify gaps in knowledge on the subject, as well as to identify areas that require more focus (Kortum, Leka and Cox, 2011; Cassar, et al., 2020) by adopting long-term solutions and policies that are data driven.

The current study aims to reconsider specific assumptions about the phenomenon of work stress and move away from simple linear assumptions to more complex understandings. Our broad research questions are therefore two: Does work related stress relate with outcomes in non-continuous ways? If so, what do such configurations tell us new about the phenomenon of stress at work? In line with the call of a number of other

scholars (e.g. Furnari, et al., 2021) we aim to sensitize management and organizational scholars to develop a sense of configurational theorizing which aims, besides other things, to look at the whole and the interdependence between constructs besides their relationships (Bacharach, 1989). We thus try and demonstrate that enhanced ways of analyses can shift assumed understandings and provide new avenues for better management practices (Cohen, 2007). We argue that this requires a degree of theoretical malleability when allowing ourselves to examine 'established' relations from a novel methodological and analytical perspective which may ultimately imply certain changed practices and interventions. For this purpose, we illustrate the potential use of fuzzy-set Qualitative Comparative Analysis (fQCA) as a specific technique using work stress-outcome relationships as an example. It is hoped that researchers come to appreciate the value of such research tools in a wider variety of business research phenomena. The rest of the review first presents a rationale for the need to approach theory and research from a critical stance in informing management practice and then we apply this rationale to the subject of work-related stress by employing fuzzy-set logic as a means to reconceptualise stress-outcome relations rather than the more typical 'linear'-based approach using correlational analyses.

2. Research as reinforcing mortar between theory and practice

There is a fine interplay between theory, research and practice with theoretical developments evolving through the re-evaluation of data through new research paradigms (Kipping and Usdiken, 2014) which then finds itself into the 'real world' as *recommendable* practices. This process is a very lengthy one because very often, these new insights, derived from more innovative research positions, are not easily accepted in the academic community. As Grey (2004, p. 178) rightly states in the case of new critical thought in management education, this may often spur "alternatives which may initially seem as outrageous as they are unfamiliar". Likewise, Birkinshaw, et al., (2014) recommend that the future of management science and organizational behaviour is to include more intriguing questions compared to the conventional attitude in management research by primarily engaging in modes of research that are both intellectually challenging and have the potential of making an impact on practice. Indeed, theory built on sound research principles translates into more effective practice. This is not a new calling though and a number of scholars have appealed for research and practice to get closer (Cohen, 2007; Latham, 2007; Rousseau, 2007; Corley and Gioia, 2011). One implication of this union would be to re-think some of the more established 'givens' in any field of business study. It implies that it may require a stronger dialogue between the interested parties and a re-evaluation of the methods for doing research, providing explanations and applying solutions (see e.g. Billig, 2013). One benefit of this is that it may prevent 'established' theories and research from mimicking each other by reconsidering the more reductionist and perhaps simplistic ways to conceive phenomena resulting in less-than-precise models. In this, the paradox is that theories may lose some of their simplicity which is what they are set to do in the first place. In fact, it would encourage departing from primarily reductionist paradigms as these rely on an underlying understanding that finding and isolating all the parts will lead to the total or sum of knowledge about a phenomenon providing predictions and replicable interventions (Bloch, 2005).

One clarification is warranted though. We must not throw the baby out with the bathwater and discard established lines of thought that have formed and moulded a specific body of knowledge, including work stress; indeed, this method of doing science has resulted in many important discoveries (Bloch, 2005). Rather it is to take a different path and use new and interesting research methodological approaches to 'build theory' besides to 'test theory' (Colquitt and Zapata-Phelan, 2007). It also means that researchers view phenomena from more complex lenses determined by non-linear dynamics to get a more precise representation of the unfolding process underlying a phenomenon. This in turn provides more reliable interventions and more insightful implications for practice. In this regard we may position or view our mode of thinking, including in the current investigation which we present as an example, as somewhere between *qualifiers* and *builders* (Colquitt and Zapata-Phelan, 2007). Thus, research strategies require a departure from the notion of finding something or nothing to ameliorate theoretical refinement (Edwards, 2010), precision, and boundaries (Edwards and Berry, 2010). The downside to presenting such a stand is of course that established theoretical frameworks are not always adequate to provide solid explanatory backdrops for complex relations because as noted by Furnari, et al., (2021), complex explanations require theories to account for simultaneous interdependencies and generally theories aim for simplicity to, ironically, avoid complexity. Hence, such perspectives may often be binned simply because they do not 'comply' with the notions of traditionally 'good theory' (Bacharach, 1989). Moreover, reinforcing this conventional position is the fact that methodological approaches and many of the adopted

techniques consider a strong direct-effect assumption thus reinforcing the shape of the theoretical phenomenon of interest.

3. The case for work stress research

Stress at work is a very well researched phenomenon, rich in research, abundant in theories and with important contributions to practice. D'Amato and Zijlstra (2003) reviewed a 10-year period of research relating occupational stress to mental health, whereas Väänänen, Murray, and Kuokkanen (2014) linked the occupational health debate to the broader historical and cultural processes that took place in organizations over a 50-year period since 1960. More recently, Cassar, et al., (2020) attempted to capture work stress research by exploring the way the field has developed in terms of its key theoretical and conceptual emphases throughout the last two decades. Such a bird's eye view fosters closer discussions on areas that may be neglected and yet are of relevance to research for practical reasons. Moreover, the field has been well covered through its many meta-analyses, which are theme-specific by virtue of their scope to test hypothesized relationships. The theme-focused meta-analyses have expectedly focused attention on specific aspects and provided a good retrospective analysis of "overall" findings and general patterns. Subjects have varied and have covered role and stress (Jackson and Schuler, 1985), gender differences (Martocchio and O'Leary, 1989), social support in stress at work (Viswesvaran, Sanchez, and Fisher, 1999), self-esteem and emotional stability (Judge and Bono, 2001), job insecurity (Sverke, Hellgren, and Näswall, 2002), job satisfaction and health (Faragher, Cass, and Cooper, 2005), work stress and Coronary Heart Diseases (CHD) (Kivimäki, et al., 2006), harassment at work (Bowling and Beehr, 2006), withdrawal behaviors (Podsakoff, LePine, and LePine, 2007), work-family conflict and satisfaction (Ford, Heinen, and Langkamer, 2007), stress management interventions (Richardson and Rothstein, 2008), and organizational constraints on employee stress (Pindek and Spector, 2016), to mention but a few. The huge majority of these investigations draw on the line that work related stressors impact outcomes and by and large conceptualise these associations within a simple theoretical frame of mind strongly emphasising a correlational-type of association and in line with the general assumptions characterising many of the theories of work stress (see Cooper, 1998) that portray primarily direct effects from stressor to outcome such as JD-R (Demerouti, et al., 2001), the challenge hindrance model (LePine, Podsakoff, and LePine, 2005) and the allostatic load model of the stress process (Ganster and Rosen, 2013).

We can comfortably argue that many of these investigations have adopted the principle of 'linearity' in line with more conventional thinking on the mechanics of work-related stress. In other words, researchers have assumed that by identifying stressors they can predict the level of strains (i.e. stress-induced outcomes) that stressors elicit. Indeed, this research agenda on work-related stress has undoubtedly generated a wealth of knowledge about the stress process and how it affects people spilling over into a number of interventions and ultimately guiding a number of established psychosocial health and safety frameworks. However, it is also known that stressful demands and outcomes are not necessarily continuously and linearly related (Frese and Zapf, 1988). Indeed, simple linear relationships are not simple at all and may mask rather more complex associations. Moreover, straightforward linear relationships are not so evident in the relationship between work-related stress and other outcomes such as performance (Edwards, Guppy and Cockerton, 2007), illness (Cohen and Manuck, 1995) and absence (Harrison and Martocchio, 1998). On the other hand, the adoption of more complex analyses has not been as common as more linearly-assuming analyses in the field of work stress (e.g. Stanton, et al., 2001). From a practical point of view, this hints at the fact that the onset of work stress may not always immediately lead to reports of ailments and that the relationship may manifest itself in a more complex and unexpected manner. Such patterns of relationships will not discard established explanatory models of work stress but may actually broaden our way of thinking between stress and outcomes and thus urge practitioners to embrace the notion that psychosocial hazards and the consequences on salient outcomes are often not as straight-forward as they may look; and this has bearing on management interventions.

4. Adopting a complex approach to stressor-outcome relationships through fuzzy sets

At its most basic form, complexity is characterised by non-linearity. Navarro, Rueff-Lopes and Rico (2020, p. 477) define non-linearity as the non-proportional relationship among variables and argue that "a straight-line representing relationships among variables will poorly describe their true linkages. Instead of looking for proportionality among variables, non-linear relations are characterised by threshold values, tipping points and continuous up and downs, among others". Adopting the use of non-linear analytical methods has huge advantages in the advancement of our understanding of organizational phenomena and this exploratory investigation examines these in line with recent calls (Navarro, Rueff-Lopes and Rico, 2020). However,

complexity also goes one step further and examines the instantaneous arrangements of variable pairs at points in the pattern of data. In line with configurational theorizing of phenomena (Furnari, et al., 2021), these approaches provide an understanding of when and how variables interrelate thus eliciting a more complete picture of the underlying dynamics inherent in the variables. From a practical point of view this also means that the onset of work-related stress may not always immediately lead to reports of ailments and that the relationship may manifest itself in a more complex manner thus providing more practical insights into the management of stress at work.

In this specific study we attempt to zoom further into idiosyncratic shifts in the relationship between stress and outcomes using a complex approach in the hope that this will urge researchers to adopt these methods more in their research on organizational phenomena in general and work-stress and outcomes, in particular. This is hoped to take the field to a fresh level of understanding that has by and large been examined using linear approaches. In order to explore these non-conventional patterns, we apply fsQCA (Ragin, 2000, 2008) to analyse specific configurations at different degrees (high/low) of perceived work-related stress on outcomes. In the area of work stress, fuzzy-sets have not been popular with very few exceptions (e.g. Chen, Jung and Peacock, 1994; Guedes, Goncalves and da Conceico Goncalves, 2017; Javed and Bartool, 2020) and mostly in relation to ergonomics and emotional labour. Indeed, stressor-outcome relations have been generally investigated using more traditional linear ways (e.g. regression analyses and SEM) and this is one of the first studies adopting this approach known to the authors.

4.1 The principles of fuzzy sets

Fuzzy-sets are based on three pillars: non-linearity, equifinality, and asymmetric associations. First, non-linearity here means that considering the same outcome, two variables can be positively related under one configuration, negatively related under another, or even unrelated. For example, in situations of high X, two other variables (Y, Z) can have different relations under different configurations. Second, is equifinality which means that there are different configurations associated with an outcome. Thus, there can be different combinations of the outcomes, assuming high and low levels of X. For example, one might have a combination of low A and low B associated with high C. At the same time, there could be another combination whereby high A and high D are associated with high X. And third, is asymmetric associations which means that what leads to the low presence of an outcome is not necessarily the inverse of what leads to the high presence of an outcome (Ragin, 2000, 2008). This implies that low Y could be related to high X across most of the configurations, but it does not exclude the fact that this outcome could also be related to X.

But what justifies this approach compared to other more traditional approaches? fsQCA is based on set-theoretic methodology and it is more justified where complex relations are presumed. Fuzzy sets approaches need to endorse the aspects of nonlinearity, equifinality, and asymmetric reciprocal relations. These complex relations cannot be assessed using the traditional linear approach. First, linear methods do not generally allow for the analysis of configurations, equifinality, or asymmetric relations. Some methods allow for the analysis of endogenous relations (e.g. SEM) and multiple interactions (e.g. interaction effects method) but analysing beyond two-way effects with these methods becomes increasingly difficult (Fiss, 2007, 2009, 2011). Second, set-theoretic approaches allow for inductive reasoning because data is analysed by case rather than by variable (Ragin, 2008). The methods focus on whether these conditions are necessary or complex. A necessary condition is a condition that has to be present for the outcome to be present. A sufficient condition is a condition associated with the outcome, but there may be other conditions. This kind of analysis also allows for a more detailed inductive reasoning (Fiss, 2007) approach in re-defining a more adequate explanation (Colquitt and Zapata-Phelan, 2007). Moreover, set-theoretic methods can be either based on crisp-set or fuzzy-set approaches. While the former approach simply indicates whether a variable is in or out of a set, the latter approach is more sophisticated as it indicates the extent by which the variable is *part* of the set. This is due to the fact that each case under the fuzzy-set approach has different degrees of membership. Such approaches are considered the most relevant when assessing complex relations (Campbell, Sirmon and Schijven, 2013; Fiss, 2007; Fiss, Cambre' and Marx, 2013; Schneider and Wagemann, 2012).

5. Method

We examined the principles of fuzz-set logic by re-running a series of correlational analyses between work related stress and salient outcomes. In doing so, we provide an example of how looking at the patterns of data

characterising the phenomenon may prime new insights into the mechanics determining the relationship between stress and outcomes and subsequently on improved management interventions.

5.1 Participants and procedures

This study was part of a wider campaign aimed at promoting psychosocial wellness at work and preventing work stress. Between March and June 2019, 509 completed questionnaires were received from employees hailing from a wide spectrum of job levels who volunteered to participate. The majority of the respondents were female (51.9%), aged 25-49 (57.8%), in full-time employment (86.8%) and in possession of an indefinite work contract (70.5%).

5.2 Measures

Work stress: One tool that has received much attention in investigating stress at work is the Health and Safety Executive's Management Standards Indicator Tool (MSIT). This consists of seven psychosocial indicators, namely demands, control, support (peer and managerial), relationships, role and change. Following the 35-item version (Edwards et al., 2008), a refined version consisting of 25 items was developed (Edwards and Webster, 2012). The MSIT's 25-item measure (Edwards, et al., 2008) received some attention in terms of its internal psychometric properties (Balducci, et al., 2017). A more recent evaluation of the shorter version (Cassar, Bezzina and Buttigieg, 2020) reported that the 25-item version had convergent and discriminant validity issues. Following the removal of five items, the revised 20-item MSIT demonstrated better construct validity and was well-represented by a higher order factor namely work-related stress (WRS). The MSIT has been shown to relate with known stress-related outcomes like general mental health, reduced job performance, intention to quit and job satisfaction (Kazi and Haslam, 2013; Marcatto, et al., 2014). Overall, these studies suggest that the modified indicator is effective in assessing psychosocial factors in stress management interventions and is a reliable assessment of WRS. Higher scores reflected higher WRS. (Cronbach's alpha = .85).

General Health: The 12-item general health questionnaire (GHQ-12) devised by Goldberg (1978) was utilised. It comprises social dysfunction, anxiety/depression and loss of confidence and item scoring ranges from '0 = not at all' to '3 = much more than usual'. An example item is: 'I have lost much sleep over worry'. Higher scores indicate lower general health/well-being. (Cronbach's alpha = .88).

Physical symptoms: The Physical Symptoms Inventory by Spector and Jex (1998) was used. It assesses physical, somatic health symptoms thought by stress researchers to be associated with psychological distress. This scale comprises 18 symptoms and respondents were required to indicate whether that had any symptoms (e.g. an upset stomach) during the past 30 days on the following scale: '0 = No', '1 = Yes, but I didn't see a doctor' and '2 = Yes and I saw a doctor'. Higher scores indicate more physical symptoms. (Cronbach's alpha = .76).

Presenteeism: The Stanford Presenteeism Scale (Koopman, et al., 2002) was utilised. It encompasses two dimensions (completing work and avoiding distractions). Item scoring ranges from '1 = strongly disagree' to '5 = strongly agree'. An example item is: "I felt hopeless about finishing certain work tasks due to my health problem/s". In this study, higher scores are indicative of higher presenteeism (i.e. a worker's inability to concentrate and accomplish work due to health problems). (Cronbach's alpha = .70).

Intention to leave: The three items pertaining to this construct were taken from Cammann, et al., (1979). Respondents rated the items on a Likert scale ranging from '1=strongly disagree' to '5= strongly agree'. An example item is: 'I often think about leaving my current job'. Higher scores are indicative of a higher intention to quit the current job. (Cronbach's alpha = .89).

All variables correlated in the proper theoretical direction (Table 1).

Table 1: Correlations between study variables

Variable	1	2	3	4	5
1. WRS	1.00	0.23**	0.17**	0.10*	0.37**
2. General health		1.00	0.49**	0.22**	0.28**
3. Physical symptoms			1.00	0.31**	0.21**
4. Presenteeism				1.00	0.12**
5. Intention to quit job					1.00

* $p < .05$; ** $p < .01$

5.3 Calibrations

One stage of analysis in fuzzy logic involves the calibration of all measures. This step involved the conversion of continuous variables to a scale which shows the strength of the variables. The scale ranged between 0 (completely out of the set) and 1 (completely in the set). The direct method of calibration is used whereby the theoretical understanding of the measures, together with the statistical distribution of each measure, are used to determine the calibrations outlined in Table 1 together with the descriptive statistics (Ragin, 2000, 2008). Before identifying the calibrations shown in Table 2, different levels of calibrations were used. That is, numerous options have been analysed (+/-5%) using variations in calibrations for each variable in order to ensure that the solutions are robust. There was no real change in the end results, meaning the model is robust.

Table 2: Descriptive statistics and calibrations

Variable	Statistical Distributions				Calibration		
	Max.	Mean	Min.	SD	Full-Membership	Cross-Over	Non-Membership
Work-related stress	4.35	2.35	1.20	0.55	4.30	2.30	1.20
General Health	2.83	0.96	0.00	0.51	2.00	0.95	0.00
Physical Symptoms	1.61	0.31	0.00	0.22	1.00	0.25	0.00
Presenteeism	5.00	2.56	1.00	0.73	4.00	2.50	1.00
Intention to Quit Job	5.00	2.44	1.00	1.24	5.00	2.20	1.00

5.4 Data analysis

The data analysis was performed using fsQCA 2.0. The data analysis involves assessing the sufficient and necessary conditions through the sufficiency analysis and necessity analysis, respectively. The former type of analysis shows different configurations that are sufficient for the outcome to occur. The latter type of analysis indicates whether a variable has to be present in order for the outcome to take place (Ragin, 2000). The outcomes for both types of analysis include high and low WRS. For both cases, the intermediate solution has been used for interpretation, as with previous studies that have adopted this approach (e.g., Moreno, Prado-Gascó, Hervás, Núñez-Pomar, and Sanz, 2016). The analysis involves a total of 4 conditions and therefore 16 (2^4) potential combinations, based on two different states, that is, high or low. The frequency threshold for the study was 10 cases per solution for both high and low outcomes, encompassing 100% of the sample, beyond the recommended 80% stipulated by Ragin (2008). In both types of analysis, in order to ensure robust results, data is examined for consistency to the proportion of cases associated with a given outcome (Ragin, 2000, 2008). Results are also examined for raw and unique coverage which “indicates how much of the membership in the outcome is covered by the membership in a single path” (Schneider and Wagemann, 2012; p. 139). Following the sufficiency analysis, the potential presence of necessary conditions is examined for both outcomes. Necessary conditions illustrate whether the condition tested has to be present for the outcome to be present (Ragin, 2000). A condition is necessary if consistency and coverage are at least 0.90 and 0.80, respectively (Ragin, 2006).

Owing to space, interested readers may find technicalities of the methods adopted in Ragin (2000).

6. Results

6.1 Necessity Conditions Results

This section will examine the necessary conditions and sufficient conditions results before presenting the associations for both high and low levels of WRS. If an outcome has a consistency and coverage of at least 0.90

and 0.80, respectively, then it implies that that specific outcome is a constant outcome for WRS. Table 3 provides a summary of the necessity conditions results.

Table 3: Necessity conditions analyses

Outcome Levels	High Presence of WRS		Low Presence of WRS	
	Consistency	Coverage	Consistency	Coverage
High General Mental Health	0.75	0.75	0.65	0.72
Low General Mental Health	0.72	0.65	0.77	0.78
High Physical Symptoms	0.73	0.74	0.65	0.74
Low Physical Symptoms	0.74	0.65	0.77	0.76
High Presenteeism	0.80	0.72	0.69	0.69
Low Presenteeism	0.65	0.65	0.72	0.80
High Intention to Leave	0.72	0.73	0.56	0.63
Low Intention to Leave	0.64	0.56	0.77	0.76

Based on the criteria identified in the methods section, none of the conditions can be considered to be necessary.

6.2 Sufficiency Conditions Results

The sufficient conditions are presented in Table 4. The table indicates bundles of outcomes that are associated with WRS. As shown in the necessity analysis, no single variable alone within the specific configurations is associated with WRS. However, as indicated in the sufficiency analysis, different bundles of outcomes are likely to be related to WRS. “●” and “⊗” illustrate the high presence and low presence of the conditions, while the blank spaces show indifference towards the outcome. We found six configurations associated with the high presence of WRS (configurations 1-6), and four configurations associated with low WRS (configurations 7-10) (Table 5). All configurations associated with high WRS share a common element, that is the high intention to leave. In configurations associated with low WRS, one can notice the absence of high intention to leave and presenteeism. The results for both high and low levels of the outcomes are explained in the sufficiency analysis below. The consistency value has been set at 0.9855 and 0.970 for the high and low outcomes, respectively, both higher than the 0.75 threshold (Ragin, 2008). Unlike consistency values, there is no particular threshold for the level of coverage and the levels of coverage are in line with studies that have adopted this method (e.g., Meuer, 2017). Overall, the solution shows clear signs of complex relations involving nonlinearity, equifinality and asymmetric relations. Nonlinearity is evident through the fact that different variables have different relations under the same outcome. Equifinality is evident through the different configurations illustrated under the two outcomes. Therefore, there are different routes associated with the two outcomes. The solution also shows asymmetric associations, when comparing the two different outcomes. This is due to the fact that configurations associated with low WRS and not the exact inverse of configurations associated with high WRS. For example, configurations 3 and 8 involve high general health and physical symptoms, yet the outcomes are different. That is configuration 3 is related to the high presence of WRS while configuration 8 is related to the low presence of WRS. These complex relations clearly show that if we had to adopt linear regression methods, the relations involved would be either over- or under estimated. These complex associations are further explained in the discussion section.

Table 4: Sufficiency conditions analyses

Permutation	High Presence of WRS						Low Presence of WRS			
	1	2	3	4	5	6	7	8	9	10
General Health	●		●		⊗	⊗		●	⊗	⊗
Physical Symptoms		⊗	●	●	⊗			●	⊗	⊗
Presenteeism	●	●		⊗		⊗	⊗		⊗	
Intention to Quit Job	●	●	●	●	●	●	⊗	⊗		⊗
Consistency	0.896	0.904	0.899	0.903	0.904	0.909	0.888	0.947	0.949	0.943
Raw Coverage	0.456	0.485	0.395	0.417	0.429	0.425	0.589	0.436	0.456	0.448
Unique Coverage	0.002	0.001	0.001	0.002	0.001	0.002	0.117	0.023	0.063	0.034
		High Presence of WRS					Low Presence of WRS			
Overall Solution Consistency		0.855					0.870			
Overall Solution Coverage		0.619					0.731			

Notes: Black circles ("●") indicate that the presence of the condition is high, and open circles ("⊗") indicate that the presence of the condition is low. Blank spaces indicate irrelevance of the condition to the solution.

Table 5: Configurations with High and Low Work-Related Stress (WRS)

WRS Presence	Configuration	Description
High	1	The person's high intention to leave is associated with high general health and presenteeism.
	2	The person's high intention to leave is associated with high presenteeism.
	3	The person's high intention to leave is associated with high general health and physical symptoms.
	4	The person's high intention to leave is associated with high physical symptoms.
	5	The person's high intention to leave is not associated with the other variables.
	6	The person's high intention to leave is not associated with the other variables.
Low	7	Low presence of variables.
	8	The person's high general health is associated with high physical symptoms.
	9	Low presence of variables.
	10	Low presence of variables.

7. Discussion

In line with our first research question, our case for stress-outcome relationships has shown that using fuzzy sets reveals that familiar associations between work stress and outcomes are not necessarily linear but rather configure themselves in more complex and indeed, more interesting ways (Furnari, et al., 2021). Indeed, associations between stress and outcomes should not always be assumed as linear or independently simple but rather that associations configure differently at different levels of high versus low levels of work-related stress. These stress-outcome relations are in line with previous suggestions and findings (Edwards, Guppy and Cockerton, 2007; Cohen and Manuck, 1995) and provide an opportunity to refine our theoretical affirmations about the phenomenon (Colquitta and Zapata-Phelan, 2007), possibly challenging conventional understanding on the mechanisms underlying the stress process. Researchers should strive to adopt such non-linear methodological approaches not to discard accepted and established knowledge but rather to refine and seek details that may ameliorate our understanding of the stress process (Edwards, 2010; Javed and Bartool, 2002; Navarro, Rueff-Lopes and Rico, 2020). In addition, results suggest that departing from a continuous approach means that interventions aimed at reducing debilitating outcomes are not likely to gradually and predictably deflate the impact of the undesirable effects in a consistent manner (our second research question). Indeed, the literature has been cautioning about this (Frese and Zapf, 1988) and it has been rehashed in more recent literature concerning the seemingly ineffectiveness of stress management interventions (Karanika-Murray and Biron, 2015). It is time researchers and practitioners give this heed. Our results do suggest that the stress process is in fact often inconspicuous.

While correlations (and regression models) with the selected outcomes included here were in the right theoretical direction and in line with the literature, the configurations underlying the structure of the associations were not constantly in the presumed linear pattern. Results suggest that the presence of high / low stress is characterized by equifinality (for example, different configurations related to specific outcomes) and causal asymmetry (for example the configuration leading to the presence of specific outcomes was not necessarily the inverse of those configurations related to its absence). For example, as indicated in Table 4, high versus low levels of work stress elicited different configurations with outcomes such that more distinct patterns emerged allowing a more detailed understanding of fine changes in the levels of stress patterns. These complex associations may also interplay as an influence of the other outcomes too that relate to each other in a somewhat fuzzy way. In addition to these complexities, the asymmetric associations in our study indicate that traditional theoretical interpretations could lead to over estimation or under estimation of the results given that the four outcomes associated to stress cannot be all equally related to WRS in a simple linear fashion. In other words, the impact of high WRS on outcomes is not the exact inverse of the impact of low WRS on that outcome. For example, high WRS is likely to be associated with high intention to leave or low levels of general mental health. Moreover, as Table 5 clearly demonstrates, high versus low levels of stress induced different configuration patterns between the outcomes themselves. This is a clean departure from our general assumptions that 'all' outcomes correlate in the same direction at all points along the continuum. Upon looking closer, the configurational analysis shows that in situations of high WRS, poor general mental health is associated with high intention to leave, while in situations of low WRS, poor general mental health is present but associated with low physical symptoms accompanied by either low presenteeism or low intention to leave. This is not at all

times though as configurations 7, 9 and 10 indicate warranting revisits to established theoretical models of WRS and outcomes and indicating that the relationship structures are slightly more complex than usually thought to be the case. The theoretical underpinnings of these findings are twofold: First, work stress-outcome relations may vary across points in the sample suggesting that participants residing in the different configurations may relate differently to different bundles of outcomes. Second, our understanding of the stress process may require us to examine further the mechanisms that generate these more granular configurations such as personality traits and other mediators.

As highlighted in section 2, these findings also have spill over implications on practice too and adopting improved or more sensitive methodological approaches refines the bridge between theory and practice (Corley and Gioia, 2011). In essence, practitioners adopting typical diagnostic tools to evaluate the impact of work stress on specific outcomes ought to keep in mind that in adopting policies and practices to reduce the impact of stressful hazards, they should equally appreciate that interventions may operate in specific ways and at specific levels and may require a combination of interventions to maintain the low level impact of specific factors on the combined effect of outcomes. For instance, improving quality services provided by Employee Assistance Programmes in high stressful circumstances may require more focus on retaining employees but in low stressful situations focus may turn on sustaining mental health. The general understanding has been that only under conditions of high stress is retention and mental health low. Fuzzy-sets approaches challenge this position. Indeed, most interventions are often a blanket-cover approach (e.g. introducing aspects in job redesign in particular work tasks) that may not be seemingly sustainable. Therefore, policy and interventions should account for this complexity. In addition, practitioners should begin analysing stress data they obtain from their respective organizations by looking at potentially more non-linear associations and assess their internal configurational patterns. This may serve to enable their future actions to generate more impactful results as they take more informed decisions.

Of course, like all studies, no investigation is without its share of limitations. In our case, the outcome variables used are not exhaustive and hence more complex pattern associations with other outcomes could be tested in future research. However, the issues raised in the stress literature and our initial results do elicit an informed confidence that a similar pattern is plausible. This should trigger a motivation to expand our mental repertoire on the mechanics underlying the stress process. Secondly, this was a cross-sectional investigation and we may find that the patterns of associations between stress and outcomes may be even more different if explored over time.

8. Conclusion

This study, using work stress-outcome relationships as an example, has demonstrated the exciting avenues presented by fsQCA. Such techniques offer a new window for researchers to explore known phenomena from a fresh perspective adding to knowledge both in terms of explanation (theory) and applications (practice). While more traditional correlational approaches should provide us with a general pattern of the relationships between variables we argue that deepening our understanding requires looking at that pattern with a more detailed eye especially if we are to attain theoretical precision and practical validity. Fuzzy-sets approaches can help us in this direction. Indeed, researchers are encouraged to deepen their knowledge about such techniques in their quest to understand finer intricacies of business research phenomena as is the case of stress at work.

References

- Bacharach, S.B., 1989. Organizational theories: some criteria for evaluation. *Academy of management review*, 14(4), pp. 496-515.
- Balducci, C., Romeo, L., Brondino, M., Lazzarini, G., Benedetti, F., Toderi, S., ... Pasini, M., 2017. The validity of the short UK health and safety executive stress indicator tool for the assessment of the psychosocial work environment in Italy. *European Journal of Psychological Assessment*, 33(3), pp. 149-157.
- Billig, M., 2013. *Learn to Write Badly: how to Succeed in the Social Sciences*. Cambridge, UK: Cambridge University Press.
- Birkinshaw, J., Healey, M.P., Suddaby, R. and Weber, K., 2014. Debating the future of management research. *Journal of Management Studies*, 51(1), pp. 38-55.
- Bloch, D.P., 2005. Complexity, chaos, and nonlinear dynamics: a new perspective on career development theory. *The career development quarterly*, 53(3), pp. 194-207.
- Bowling, N. A. and Beehr, T. A., 2006. Workplace harassment from the victim's perspective: a theoretical model and meta-analysis. *Journal of Applied Psychology*, 91, pp. 998-1012.
- Cammann, C., Fichman, M., Jenkins, D. and Klesh, J., 1979. The Michigan Organizational Assessment Questionnaire. Unpublished manuscript, University of Michigan, Ann Arbor.

- Campbell, J. T., Sirmon, D. and Schijven, M., 2013. *The good, the bad, and the fuzzy: a configurational approach to acquisitions*. Paper presented at the seventy-third annual meeting of the Academy of Management, Orlando, FL.
- Cassar, V., Bezzina, F. and Buttigieg, S., 2020. Investigating the psychometric properties and assessment capabilities of the short version of the Health and Safety Executive's Management Standards Indicator Tool. *The International Journal of Human Resource Management*, 31(16), pp.2115-2140.
- Cassar, V., Bezzina, F., Fabri, S. and Buttigieg, S.C., 2020. Work stress in the 21st century: a bibliometric scan of the first 2 decades of research in this millennium. *The Psychologist-Manager Journal*, 23(2), pp. 47-75
- Chen, J.-G., Jung, H. S., & Peacock, B. J., 1994. A fuzzy sets modelling approach for ergonomic workload stress analysis. *International Journal of Industrial Ergonomics*, 13(3), pp. 189-216.
- Cohen, D.J., 2007. The very separate worlds of academic and practitioner publications in human resource management: reasons for the divide and concrete solutions for bridging the gap. *Academy of Management Journal*, 50(5), pp. 1013-1019.
- Cohen, S. and Manuck, S. B. 1995. Stress, reactivity and disease. *Psychosomatic Medicine*, 57, pp. 423-426.
- Colquitt, J.A. and Zapata-Phelan, C.P., 2007. Trends in theory building and theory testing: A five-decade study of the Academy of Management Journal. *Academy of management journal*, 50(6), pp. 1281-1303.
- Cooper, C.L. ed., 1998. Theories of organizational stress. Oxford: Oxford University Press.
- Corley, K.G. and Gioia, D.A., 2011. Building theory about theory building: what constitutes a theoretical contribution?. *Academy of management review*, 36(1), pp. 12-32.
- Cox, T., Griffiths, A. and Rial-Gonzalez, E., 2000. Work-related stress. *Office for Official Publications of the European Communities, Luxembourg*.
- D'Amato, A. and Zijlstra, F. R. H., 2003. *Occupational stress: A review of the literature relating to mental health. Stress Impact Surrey*. Guildford, United Kingdom: University of Surrey.
- Demerouti, E., Bakker, A. B., Nachreiner, F. and Schaufeli, W. B., 2001. The job demands-resourcesmodel of burnout. *Journal of Applied Psychology*, 86, pp. 499-512.
- Edwards, J.R., 2010. Reconsidering theoretical progress in organizational and management research. *Organizational Research Methods*, 13(4), pp. 615-619.
- Edwards, J.R. and Berry, J.W., 2010. The presence of something or the absence of nothing: Increasing theoretical precision in management research. *Organizational Research Methods*, 13(4), pp. 668-689.
- Edwards, J. and Webster, S., 2012. Psychosocial risk assessment: Measurement invariance of the UK Health and Safety Executive's Management Standards Indicator Tool across public and private sector organizations. *Work & Stress*, 26, pp. 130-142.
- Edwards, J.A., Guppy, A. and Cockerton, T., 2007. A longitudinal study exploring the relationships between occupational stressors, non-work stressors, and work performance. *Work & Stress*, 21(2), pp .99-116.
- Edwards, J., Webster, S., Van Laar, D. and Easton, S., 2008. Psychometric analysis of the UK Health and Safety Executive's Management Standards work-related stress Indicator Tool. *Work & Stress*, 22, pp.96-107.
- Faragher, E. B., Cass, M. and Cooper, C. L., 2005. The relationship between job satisfaction and health: a meta-analysis. *Occupational and Environmental Medicine*, 62, pp. 105-112.
- Fiss, P. C., 2007. A set-theoretic approach to organizational configurations. *Academy of Management Review*, 32(4), pp.1180-1198.
- Fiss, P. C., 2009. Case studies and the configurational analysis of organizational phenomena. In: C. R. Charles & D. Byrne, eds. 2009. *The Handbook of Case Study Methods*. London, U.K.: Sage. pp.424-440.
- Fiss, P. C., 2011. Building better causal theories: a fuzzy set approach to typologies in organization research. *Academy of Management Journal*, 54(2), pp. 393-420.
- Fiss, P. C., Cambré, B. and Marx, A., 2013. *Configurational Theory and Methods in Organizational Research*. Bingley, U.K.: Emerald Group Publishing.
- Ford, M. T., Heinen, B. A. and Langkamer, K. L., 2007. Work and family satisfaction and conflict: a meta-analysis of cross-domain relations. *Journal of Applied Psychology*, 92, pp. 57-80.
- Furnari, S., Crilly, D., Misangyi, V.F., Greckhamer, T., Fiss, P.C. and Aguilera, R., 2021. Capturing causal complexity: heuristics for configurational theorizing. *Academy of Management Review*, 46(4), pp. 778-799.
- Ganster, D. C. and Rosen, C. C., 2013. Work stress and employee health: a multidisciplinary review. *Journal of Management*, 39, pp.1085-1122.
- Goldberg, D., 1978. *Manual of the GHQ*. Windsor: NFER-Nelson.
- Grey, C., 2004. Reinventing business schools: the contribution of critical management education. *Academy of Management Learning & Education*, 3(2), pp. 178-186.
- Guedes, M. J., Gonçalves, H. M. and da Conceição Gonçalves, V., 2017. Stress at the top: myth or fact? causal explanations from a fuzzy-set qualitative comparative analysis (fsQCA). *Quality & Quantity*, 51(5), pp. 2001-2017.
- Hassard, J., Teoh, K.R., Visockaite, G., Dewe, P. and Cox, T., 2018. The cost of work-related stress to society: A systematic review. *Journal of occupational health psychology*, 23(1), pp. 1-17.
- Jackson, S. E. and Schuler, R. S., 1985. A meta-analysis and conceptual critique of research on role ambiguity and role conflict in work settings. *Organizational Behavior and Human Decision Processes*, 36, pp. 16-78.
- Javed, S., and Batool, N. (2020). Assessment of conventional and advanced methodologies: Application of configurational approach and fuzzy set theory in emotional labor research, *The Electronic Journal of Business Research Methods*, 18(2), pp.72-83, available online at www.ejbrm.com.

- Judge, T. A. and Bono, J. E., 2001. A rose by any other name: Are self-esteem, generalized self-efficacy, neuroticism, and locus of control indicators of a common construct? In: B. W. Roberts & R. Hogan, eds. 2001. *Decade of behaviour: personality psychology in the workplace*. Washington, DC: American Psychological Association. pp. 93-118.
- Kazi, A. and Haslam, C.O., 2013. Stress management standards: a warning indicator for employee health. *Occupational Medicine*, 63(5), pp. 335-340.
- Kipping, M. and Üsdiken, B., 2014. History in organization and management theory: more than meets the eye. *Academy of Management Annals*, 8(1), pp. 535-588.
- Kivimäki, M., Virtanen, M., Elovainio, M., Kouvonen, A., Väänänen, A. and Vahtera, J., 2006. Work stress in the etiology of coronary heart disease: a meta-analysis. *Scandinavian Journal of Work, Environment and Health*, 32, pp. 431-442.
- Koopman, C., Pelletier, K.R., Murray, J.F., Sharda, C.E., Berger, M.L., Turpin, R.S., Hackleman, P., Gibson, P., Holmes, D.M. and Bendel, T., 2002. Stanford presenteeism scale: health status and employee productivity. *Journal of occupational and environmental medicine*, 44(1), pp. 14-20.
- Kortum, E., Leka, S. and Cox, T., 2011. Perceptions of psychosocial hazards, work-related stress and workplace priority risks in developing countries. *Journal of Occupational Health*, pp. 225-238.
- Latham, G.P., 2007. A speculative perspective on the transfer of behavioral science findings to the workplace: "the times they are a-changin'". *Academy of Management Journal*, 50(5), pp. 1027-1032.
- Leka, S., Cox, T. and Zwetsloot, G.I.J.M., 2008. The European framework for psychosocial risk management (PRIMA-EF). In: S. Leka and T. Cox, eds. 2008. *The European Framework for Psychosocial Risk Management: PRIMA-EF*. Brussels: EU Commission. pp. 1-16.
- Lepine, J. A., Podsakoff, N. P. and Lepine, M. A., 2005. A meta-analytic test of the challenge stressor–hindrance stressor framework: an explanation for inconsistent relationships among stressors and performance. *Academy of Management Journal*, 48, pp. 764-775.
- Marcatto, F., Colautti, L., Filon, L., Luis, O. and Ferrante, D., 2014. The HSE management standards indicator tool: concurrent and construct validity. *Occupational Medicine*, 64(5), pp. 365-371.
- Martocchio, J. J. and O'Leary, A. M., 1989. Sex differences in occupational stress: a meta-analytic review. *Journal of Applied Psychology*, 74, pp. 495-501.
- Moreno, F.C., Prado-Gascó, V., Hervás, J.C., Núñez-Pomar, J. and Sanz, V.A., 2016. Predicting future intentions of basketball spectators using SEM and fsQCA. *Journal of Business Research*, 69(4), pp. 1396-1400.
- Meuer, J., 2017. Exploring the complementarities within high-performance work systems: a set-theoretic analysis of UK firms. *Human Resource Management*, 56(4), pp. 651-672.
- Navarro, J., Rita Rueff-Lopes, R. and Rico, R., 2020. New nonlinear and dynamic avenues for the study of work and organizational psychology: an introduction to the special issue. *European Journal of Work and Organizational Psychology*, 29(4), pp. 477-482.
- Pindek, S. and Spector, P. E., 2016. Organizational constraints: a meta-analysis of a major stressor. *Work and Stress*, 30, pp. 7-25.
- Podsakoff, N. P., LePine, J. A. and LePine, M. A., 2007. Differential challenge stressor hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behavior: a meta-analysis. *Journal of Applied Psychology*, 92, pp. 438-454.
- Ragin, C.C., 2000. *Fuzzy-set social science*. Chicago and London: University of Chicago Press.
- Ragin, C.C., 2006. Set relations in social research: evaluating their consistency and coverage. *Political analysis*, (Summer 2006), pp. 291-310.
- Ragin, C. C., 2008. Qualitative comparative analysis using fuzzy sets (FsQCA). In: B. Rihoux and C. C. Ragin, eds. *Configurational Comparative Methods, Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks, CA: Sage. pp. 87-122.
- Richardson, K. M. and Rothstein, H. R., 2008. Effects of occupational stress management intervention programs: a meta-analysis. *Journal of Occupational Health Psychology*, 13, pp. 69-93.
- Rousseau D.M., 2007. A sticky, leveraging, and scalable strategy for high-quality connections between organizational practice and science. *Academy of Management Journal*, 50(5), pp. 1037-1042.
- Schneider, C. Q. and Wagemann, C., 2012. *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge, U.K.: Cambridge University Press.
- Spector, P.E. and Jex, S.M., 1998. Development of four self-report measures of job stressors and strain: interpersonal conflict at work scale, organizational constraints scale, quantitative workload inventory, and physical symptoms inventory. *Journal of occupational health psychology*, 3(4), pp. 356-367.
- Stanton, J.M., Balzer, W.K., Smith, P.C., Parra, L.F. and Ironson, G., 2001. A general measure of work stress: the stress in general scale. *Educational and Psychological Measurement*, 61(5), pp. 866-888.
- Sverke, M., Hellgren, J. and Näswall, K., 2002. No security: a meta-analysis and review of job insecurity and its consequences. *Journal of Occupational Health Psychology*, 7, pp. 242-264.
- Väänänen, A., Murray, M. and Kuokkanen, A., 2014. The growth and the stagnation of work stress: publication trends and scientific representations 1960-2011. *History of the Human Sciences*, 27, pp. 116-138.
- van Stolk, C., Staetsky, L., Hassan, E. and Kim, C. W., 2014. *Management of psychosocial risks at work: An analysis of the findings of the European Survey of Enterprises on New and Emerging Risks (ESENER) European Risk: A European Risk Observatory Report*. Belgium: EU-OSHA.

- Viswesvaran, C., Sanchez, J. I. and Fisher, J., 1999. The role of social support in the process of work stress: a meta-analysis. *Journal of Vocational Behavior*, 54, pp. 314–334.
- Zoni, S., & Lucchini, R., 2012. European approaches to work-related stress: a critical review on risk evaluation. *Safety and Health at Work*, 3(1), pp. 43–49.

Network Analysis of Traditional Word of Mouth

Aliosha Alexandrov¹, Sandipan Sen² and Michael J. Tippins¹

¹University of Wisconsin Oshkosh, College of Business Administration, Oshkosh, Wisconsin, USA

²Southeast Missouri State University, Cape Girardeau, Missouri, USA

alexanda@uwosh.edu

ssen@semo.edu

tippins@uwosh.edu

Abstract: Network analysis of Word of Mouth (WOM) examines how customers exchange opinions within their social networks. Compared to standard survey questions, which typically measure the likelihood to recommend, the network approach provides more metrics (e.g., average path length, clustering coefficient, density, average degree) that can be used to diagnose customer chatter. Unfortunately, traditional WOM has not benefitted from network analysis, which usually is applied to online WOM due to the availability of stored data. Despite the pervasiveness of online WOM, recent commercial reports reveal that traditional WOM still surpasses online WOM by a large margin. Traditional WOM is also perceived as more trustworthy and persuasive than online WOM. Thus, given the heavy reliance on traditional WOM and the advances made in network analysis that deal with online WOM, the main purpose of this study is to demonstrate that network analysis is a viable option for gaining insights in traditional WOM. This is the first study to utilize a survey method for collecting WOM network data for a wide range of products, which allows a direct comparison of their network structures. While network analysis may be more demanding on the researcher and the respondents, as the study illustrates, it also is more diagnostic than a standard survey. The main study utilized network analysis by using an alter-alter method, which was used to map the WOM networks structures for multiple products. Specifically, we examined the WOM networks structure as a function of product type (search, experience, and credence products) and opinion valence (positive vs. negative). The results reveal that WOM networks are affected primarily by product type. People are most likely to share opinions about experience products, followed by opinions about search products, and least likely to talk about credence products. The effect of opinion valence is limited. These findings are of practical relevance because they show that WOM can be managed by including search, experience, or credence qualities in promotional messages. This study also is the first to compare WOM networks to the existing social network, which can serve as a benchmark for evaluating WOM campaigns. The results reveal that for most products, people do not utilize all of their social connections for WOM, but there are exceptions, such as sharing a positive opinion about a movie, where WOM chatter can exceed the social network. Introducing a new method for studying traditional WOM also brings new research questions, and the study concludes with limitations and future research directions. Overall, the conclusion is that network analysis is a viable technique for studying traditional WOM.

Keywords Word of Mouth; WOM Valence, Network Analysis; Social Networks; Search, Experience, Credence

1. Introduction

Network analysis of Word of Mouth (WOM) examines how customers exchange opinions within their social networks. Compared to standard survey questions, which typically measure the likelihood to recommend, the network approach goes one step further and examines to who else in the network a person talks to. This allows capturing the paths of WOM within society, thus forming a network of opinions. The network approach provides multiple metrics of practical relevance; for example, how quickly WOM disseminates through society; to how many people, on average, a person shares WOM; what proportion of the population shares WOM; or whether people tend to share WOM within their social group, or randomly to everyone else.

Network analysis is applied mainly to online WOM. One reason for this is the availability of stored data from social media websites and emails, where the interconnectedness of people can be extracted. Traditional WOM, on the other hand, almost never benefited from network analysis. The main reason for this is that collecting personal

network data in an offline context is challenging and the methodology remains unfamiliar (McCarty, Killworth, and Rennell, 2007). This asymmetry leaves the social network aspects behind traditional WOM largely unexplored.

An indirect reason for the lack of network analyses in traditional WOM is the increased interest in online WOM. Currently, the ratio of published research on online/electronic WOM vs. offline/traditional WOM for the period 2015-2020 is: 2:1 at ABI/Inform; 1.5:1 at Google Scholar; and 7:1 at Business Source Premier. However, the increased research interest in online WOM should not negate the importance of traditional WOM. Traditional WOM surpasses online WOM and even during the COVID-19 pandemic, 83% of WOM was traditional (66% face-to-face and 17% on the phone), (Engagement Labs, 2021). Overall, practitioners advise that failing to pay attention to natural conversations is a mistake because online and offline WOM are different, but work seamlessly together (Fulgoni and Lipsman, 2015).

Considering the strong standing of traditional WOM and the advances in network analysis due to online WOM, this study fills a gap by demonstrating how a network analysis can be applied to traditional WOM to answer questions of practical relevance. The survey methodology we suggest is applicable for academic or applied research.

2. Literature Review

2.1 Traditional WOM

Research on traditional WOM is substantive and here are some selected highlights. The value of WOM is in its persuasive power. It was found fairly early that WOM is more influential than advertising (Katz and Lazarsfeld, 1955) and it has remained in the focus of marketers ever since (Berger, 2016). Customers acquired through WOM referrals have higher customer equity and generate more profit (Villanueva et al., 2008; Trusov et al., 2009; Schmitt, Skiera, and Van den Bulte, 2011). WOM plays a key role for service organizations, especially those offering complex services that are difficult to evaluate as it is perceived to be highly credible information (Silverman, 2001). From the perspective of a company, De Matos and Rossi (2008) summarize that WOM is driven by satisfaction, loyalty, quality, commitment, trust, and perceived value. From the perspective of customers, however, WOM is driven by self- and social-motives like the desire to achieve self-enhancement, self-affirmation, bond with others, altruism, helping behavior, etc. (Alexandrov et al., 2013, Berger, 2014). Customers with extreme levels of satisfaction or dissatisfaction are more likely to share experiences than customers with neutral opinions (Anderson, 1998). Communication channels (i.e., oral vs. written) affect the content of the message (Berger and Iyengar 2013), which suggests that online and traditional WOM elicit different responses.

2.2 Why Traditional WOM Matters

Compared to online WOM, there are several reasons to continue researching traditional WOM. First, there have been concerns about the trustworthiness of online WOM and the ease of its manipulation (Anderson and Simester, 2014; Aral, 2014; Luca and Zervas, 2016; Mayzlin, Dover, and Chevalier, 2014). Even B2B buyers trust referrals from friends and colleagues more than online reviews or information provided by vendors (Neumann, 2017). Amazon has a problem with fake reviews (Nguyen, 2021). In 2020, Amazon stopped more than 200 million fake reviews (Coldewey, 2021), and is seeking help from social media companies to curtail the problem (Musil, 2021). Most of the reviews on Amazon are authentic, but for some categories, the proportion of fake reviews can be as high as 61% (Sterling, 2018), which often are generated by “fake review factories” that run on Facebook (Collinson, 2018; Dwoskin and Timberg, 2018). Another practice that may affect the trust in online WOM is the growing influencer industry, which is becoming a profitable business model. The goal of influencers is to acquire paid promotions, and to achieve higher visibility, some of them purchase fake followers (Confessore et al., 2018).

Offline WOM has a stronger effect on purchase intentions than online WOM (Baker, Donthu, and Kumar, 2016). In a meta-analysis, Rosario et al. (2016) found that WOM has a stronger effect for tangible goods than for services, which suggests that not all products benefit equally from WOM. The evidence presented in traditional WOM has a stronger effect on its perceived usefulness than the evidence presented in online WOM (Martin and Lueg, 2013). Meuter, McCabe, and Curran (2013) compared traditional WOM with different forms of online WOM and concluded that traditional WOM referrals had more influence on respondents' trust perceptions and behavioral intentions about a choice of a service provider compared all other forms of online WOM. Online WOM does not predict offline

behavior well (Fay and Larkin, 2017). Finally, people are more likely to share WOM traditionally face-to-face than online (Eisingerich et al., 2015).

2.3 WOM as a Network

In general, network analysis examines the structure and behavior of complex networks. For example, the structure of the power grid (Pagani and Aiello, 2013) or the resilience of the Internet against breakdowns (Cohen et al., 2011) are topics based on network analysis. In social sciences, network analysis has been advanced mainly through sociology (e.g., Carrington, Scott, and Wasserman, 2005; Wasserman and Faust, 1994) as it examines the interconnectedness of people in the society. Social networks have a specific structure: people are densely interconnected in social cliques, and the cliques are loosely connected to each other. Network analysis has become particularly popular in social sciences after the seminal work of Watts and Strogatz, (1998), who formulated the mathematical mechanics of clustered networks, where clusters are similar to social cliques. The authors called clustered networks “small-worlds”, and it was quickly noted that social networks can be described by the mechanics of small-worlds (Watts, 1999, 2003; Newman and Park, 2003). One advantage of small-world networks is that they are more efficient in disseminating information when compared to random networks, which have no clustering but random connections (Watts, 2003). That is, for the same number of connections, it is faster to reach on average any other connection in a small-world network than in a random network.

Network structures can be represented by multiple parameters, but network analysis of online WOM has focused on a few useful parameters (Libai et al., 2013). Table 1 lists the definition of four network parameters and their interpretation in a WOM context. This study demonstrates how the network parameters in Table 1 can be estimated for different products.

Table 1: Summary of Network Properties

Network parameter	Description	WOM interpretation
Average path length (L)	The average number of hops a person needs to make to reach any other person in a network. The famous “six degrees of separation” is $L=6$ (Travers and Milgram, 1967).	On average, how many times WOM should be shared to reach any other person in a network.
Clustering coefficient (C)	The average interconnectedness of a person’s friends. If <i>none</i> of the friends of a person know one another, then $C=0$; and if <i>all</i> friends of a person know one another, then $C=1$. It is the average probability that any two friends of a person know each other (Watts, 2003; Newman, 2003a).	The average WOM chatter among the friends of a person. If <i>none</i> of the friends shares WOM among themselves, then $C=0$; and if <i>all</i> friends of a person share WOM to everyone else, then $C=1$.
Network Density	The ratio of the existing connections to all possible theoretical connections in a network.	The ratio of the actual WOM to the maximum WOM (i.e., all people talking to one another).
Average degree	The average number of social connections a person has.	The average number of people to which a person shares WOM.

Network analysis allowed for new insights in WOM, some of which are described next. Seeding a WOM campaign using the network of the well- connected people can increase the effectiveness of the campaign by eight times compared to a random seeding strategy (Hinz et al., 2011). WOM campaigns generate value in two ways: by acquiring customers who would not otherwise purchase; and by accelerating a purchase (Libai et al., 2013). Network structure affects customer defection (i.e., brand switching), and the exposure to a defecting friend increases the defecting hazard by 80% (Nitzan and Libai, 2011). In the initial phase of product adoption, the WOM networks are small, and friends have a stronger and more consistent effect than strangers, but as the network grows, the effect of strangers increases (Zhang, Liu, and Chen, 2015). Overall, social network structure is essential for achieving social contagion, even when controlled for marketing efforts (Iyengar, Van den Bulte, and Valente, 2011, Watts 2007). A simulation study by Li and Du (2017) showed that online WOM is better suited for customers with diverse opinions, but offline WOM is better suited for higher margin products.

All of the aforementioned insights, however, are based exclusively on online WOM or simulations. There only are a handful of studies examining traditional WOM as a network. Brown and Reingen (1997) examined the effect of social ties in a WOM network. Frenzen and Nakamoto (1993) were one of the first to discuss the effect of WOM structure on retransmission of opinion. More recent studies examine WOM networks but only in a healthcare context (Idonthyengar, Van den Bulte, and Valente, 2011). Even large scale studies comparing offline WOM and offline WOM examine it only on individual level without estimating macro network properties (Baker, Donthu, and Kumar, 2016). To the authors' knowledge, currently there are no studies examining the network structure of traditional WOM in a broader context or providing guidelines how it can be done or used in practice. This study fills the gap and helps keep research on traditional WOM on par with online WOM.

3. Theoretical Development

3.1 Factors Affecting WOM

People do not share WOM with everyone, but approach it selectively. For example, they may share opinions about health services with some people, and opinions about cars with others. Therefore, WOM can spread in different ways through a social network. This study explores two factors that can affect the structure of traditional WOM networks: *valence of opinions* and *product type*. WOM valence (i.e., positive vs. negative) is studied well (de Matos and Rossi, 2008; Berger, 2014; Liu et al., 2021) and is of key importance because naturally marketers want to maximize positive WOM and minimize negative WOM.

Network research for product type is limited for several reasons. Previous studies often use simulation to support their findings (Watts and Dodds, 2007; Libai et al., 2013). The analyzed data often is secondary (e.g., emails, YouTube, etc.) (De Bruyn and Lilien, 2008; Liu-Thompkins, 2012) and limited to popular online products like electronics (Smith et al., 2007; Hinz, Schulze, and Takac, 2014). Even in the large scale study of Baker, Donthu, and Kumar 2016, who utilized customer social ties in their analysis to compare online WOM and offline WOM, the impact of product-type was not considered while examining purchase and retransmission intentions. From a networking perspective, knowing how WOM for different product types is likely to spread within a social network is of high practical importance because it can help understand what type of products are fit for WOM.

At a general level, product type can be classified on the search-experience-credence continuum (Ford, Smith, and Swasy 1988; Zeithaml, 1981). In brief, products high on search qualities can be evaluated before they are purchased (e.g., laptop, jacket, etc.), products high on experience qualities can be evaluated only after they are consumed (e.g., movie, restaurant, vacation, etc.), and products high on credence qualities are difficult or impossible to evaluate even after they are consumed (e.g., healthcare, legal services, car repairs, etc.). It has been suggested that most goods are high on search and experience qualities, and most services are high on experience and credence qualities (Zeithaml, 1981).

3.2 Research Questions

The study answers three research questions regarding the effect of product type and opinion valence on the network structure of traditional WOM.

It is expected that product type, as represented by the search-experience-credence framework affects WOM. The reason for it is that the perceived risk increases moving from search and experience to credence qualities (Mitra, Reiss, and Capella, 1999). Sharing WOM may pose a risk to the sender, which Frenzen and Nakamoto (1993) call moral hazard. For example, the difficulty of evaluating credence qualities leaves a person with incomplete information and decreases their ability to help others. Also, some customer experiences may be too subjective, and thus not transferable to others. In such scenarios, customers may feel hesitant to share their opinions. According to Frenzen and Nakamoto (1993), the higher the moral hazard, the lower their willingness to share WOM. Therefore, it is important to understand the effect of product type on WOM.

RQ1: What is the effect of product type on WOM?

It has been suggested that people are more likely to share negative WOM than positive WOM (Hanna et al., 2013). The empirical evidence, however, does not confirm this. Some studies find a small, but not substantial, difference

(Anderson, 1998; East et al., 2007), and others find no statistical difference between positive and negative WOM (Jansen et al., 2009). Examining the total WOM from a large database, Baker, Donthu, and Kumar (2016) found that people share up to seven times more positive WOM than negative WOM, which was similar to the findings of Carl (2006). There is also the argument that people share negative WOM because of its diagnostic nature while spreading of positive WOM arises due to its assisting role in a decision-making process (Bi, Zhang, and Ha, 2019). None of the cited research, however, has compared positive and negative WOM using a network approach.

RQ2: What is the effect of opinion valence on WOM?

People are connected through their social networks, but they may not indulge in WOM behavior with all social ties. From a managerial point of view, it is useful to know the extent to which WOM would spread through a social network. If WOM cannot reach the capacity of the existing social network, then it may not be a useful promotional tool. Therefore, comparing WOM networks for different products types and opinions to the entire social network is essential. The social network can serve as a benchmark for evaluating the effectiveness of WOM campaigns. Understanding when WOM will spread to a large rather than a smaller fraction of the social network is critical for allocating promotional spending. To the best of the authors' knowledge, no studies have compared WOM network to a social network.

RQ3: How do WOM networks (i.e., for different products types and valence) compare to a social network?

4. Research Method

4.1 Product Selection

An initial selection including 15 products was created based on established prior research (Murray, 1991; Ford, Smith, and Swasy, 1988), and an informal focus group with the intended respondents confirmed the relevance to their lifestyle. The guiding principle for the product selection was the search-experience-credence definition by Ford, Smith, and Swasy (1988). The initial products were: for search products - *windbreaker jacket, laptop, barbeque grill, running shoes, digital camera*; for experience products - *restaurant meal, haircut, furniture rental, vacation, movie*; and for credence products - *teeth cleaning, tax advice, home maintenance, eye examination, auto repair*.

Table 2: Pretest Results for Product Type Classification

Product Type	Search-Experience-Credence Agreement	Goods vs. Services
<i>Search</i>		
Windbreaker Jacket	80%	1.08
Laptop Computer	76%	1.71
Digital Camera	78%	1.57
<i>Experience</i>		
Restaurant Meal	96%	4.37
Vacation	94%	5.63
Movie	94%	4.70
<i>Credence</i>		
Tax advice/preparation	59%	6.65
Home Maintenance	60%	6.41
Auto Repair	61%	6.39

A pretest study checked which products correctly represented the search-experience-credence classification. Forty-nine students from a Midwestern university in the United States of America were provided with the definitions of each product category, and then asked to classify the 15 products as either search, experience, or credence dimensions. Following Girard and Dion (2010), we retained products with at least 50% agreement among respondents. In addition to the product classification question, respondents were asked to evaluate on a 7-point scale the degree to which each product was a typical goods or service (1-Goods to 7-Service). The results in Table 2 are consistent with Zeithaml (1981), indicating that most search products are goods, most credence products are

services, and experience products share goods and service qualities, which confirms that the final products selection is appropriate.

4.2 Supplementary Study: Online WOM vs. Offline WOM Sharing Intentions

In a supplementary study, we tested the preferred type of WOM (i.e., online vs. traditional) for the selected products. Eisingerich et al. (2015) suggested a preference for traditional WOM over online WOM, but we wanted to verify it in the context of this study. A short survey asked how likely a person would be to share their opinion (valence was not specified) about each of the 9 products on an online platform or in a traditional conversational. Each response was measured on a *1-Not at all* to *5-Definitely will* scale. The survey was sent to random respondents on Amazon Turk, from where we obtained 89 usable responses. Each respondent answered multiple WOM questions, and to remove any interpersonal variation, we used a general linear mixed model, where respondents were coded as random variables and the averaged product types (i.e., search, experience, credence) and WOM type (i.e., online vs. traditional) were fixed factors. Based on the effect size represented by the Partial Eta Squared, η^2 , all effect were large with $\eta^2 > .14$. The highest variation in WOM was due to interpersonal variation ($\eta^2=.36$), followed by WOM type ($\eta^2=.28$) and product type ($\eta^2=.27$). All effects were significant at $p<.00$. The results confirm that, on average, people are more likely to share traditional WOM by 1.12 points on the 5-point scale when compared to online WOM. This finding confirm the preference for traditional WOM for the products in the study.

An interesting result emerged about product type, where WOM for experience products was the most likely to be shared ($b=1.33$), followed by WOM for search products ($b=.56$), and WOM for credence products ($b=.00$), which served as a benchmark in the model. Although not discussed here, the network analysis approach in the main study confirms this pattern.

4.3 Main Study Data Collection

The goal of network analysis is to map how WOM flows through society, but the data collection for networks comes with two caveats. First, traditional WOM requires sampling from a network, because unlike online WOM, face-to-face conversations are not stored in databases. In contrast to traditional statistics, however, for most network parameters, there is no inferential mechanism to generalize from a sample to the population. The estimation of a network parameter can be affected by a variety of factors including: missing data (Kossinets, 2006), the size of the sample (Anderson, Butts, and Carley, 1999), or the sampling method (Frank, 2002). Second, network analysis produces large datasets organized in matrixes. Usually, a question on a survey results in a single data column. In network analysis, each respondent answers a single question in reference to multiple other respondents in the social network, which results in a matrix (a.k.a. adjacent matrix). This data collection method is called alter-alter, where alters are the connections in a person's network. Thus, each question results in an adjacent matrix representing the network of relationships for the particular question. Therefore, one question asked to n respondents can generate up to $n(n-1)$ data points. Because the alter-alter method is demanding on respondents, the samples usually are small (McCarty, Killworth, and Rennell, 2007).

Despite the aforementioned shortcomings, previous research has established practical guidelines for studying social networks. Studies based on simulations have demonstrated that sampling still provides good parameter estimates (Galaskiewicz, 1991). Sampling about 10% of a medium size population (e.g., $N<500$) has shown to capture representative information about the social network (McCarty, Killworth, and Rennell, 2007). For large populations, Leskovec and Faloutsos (2006) recommend sampling a minimum 15% of the population. Ideally, it is recommended to sample 20%-40% of the network to guarantee accurate estimates of all network parameters. According to Iyengar, Van den Bulte, and Valente (2011), however, if the goal of the research is not to estimate the parameters of an entire network (e.g., for epidemiological research), but is focused on differences between network parameters, then using a sample is appropriate. In other words, sampling can be used for relative comparisons. Finally, one widely applied practice is to validate network estimates externally by comparing them to the properties of previously estimated networks to make sure they are in line with expectations (Libai et al., 2013).

There are probabilistic and non-probabilistic sampling techniques for networks (Leskovec and Faloutsos, 2006), two of which should be noted. One approach is to sample nodes randomly, which provides fairly good estimates. However, random sampling is not the best sampling technique. A second approach is the "forest fire" sampling

method, where the sample is a patch of the network similar to burnt forest and provides better overall estimates. Importantly, “forest fire” is *not* a form of random sampling, and captures a complete network segment without randomization. This may be counterintuitive to traditional statistical reasoning, where randomization is preferred, but in network analysis, it is more important to not lose connections. Both sampling techniques provide good estimates of average path length, clustering coefficient and other network parameters, which are of interest to this study.

Based on the guidelines above and the intention of the study, we randomly recruited 67 students, which represents 12% of all students in a college at a Midwestern university. The college, according to the participants, was their natural environment and defined their typical social network. Because we intended to collect information about multiple products, and because the alter-alter method is demanding on respondents, the data collection consisted of three rounds over a two-week period. The need for prior commitment of the respondents was necessary in order to avoid comparing WOM networks based on different people. It also allowed listing all names on the surveys prior to the data collection. Essentially, our approach amounted to a random sample, which was held constant. The reason was that it was impossible to “start a fire” for every product and keep the respondents the same for all products. In addition, it allowed comparing the WOM network for all products to a single social network serving as a benchmark. Had we focused on WOM for a single product, then we could have applied the forest fire technique, which is discussed in the implications section.

The intention to share positive and negative WOM was measured for each of the nine products, and two questions measured the strength of social ties to every person in the network. The WOM questions were dichotomous, using a “Yes/No” check format. For example: “Imagine that you had a [positive/negative] experience for a particular [product]. Would you share your experience with [student name] (spontaneously or if asked)?” The products and valence were randomized. The underlying social network was measured with the following questions: “How close is your relationship with [student name]?” ranging from *1-Not at All Close* to *7-Very Close*, and “How often do you interact with [student name]?” ranging from *1-Never* to *7-Very Often*. An example of the survey is included in Figure 1. Although providing only checkmarks, answering the 20 questions for *each* of the 67 members in the segment was daunting, and the three data collection rounds eased the burden. Six respondents did not participate in all rounds, and their responses were removed, which resulted in 20 (61 x 61) adjacent matrixes.

	Product 1	Product 2	Product 3
Name 1	x	x	x
Name 2		x	
Name 3	x		x
...			
Name 67	x	x	

Figure 1: Survey Format Example

5. Results

Examining the research questions proceeded with combining the nine products into search, experience, and credence categories for both positive and negative WOM. For example, the three experience products (restaurant meal, vacation, and movie) for positive WOM were combined into a single network. This was done by retaining only the common ties used to share information for all products within a category. For example, if a person shared their opinion about digital camera to 5 people, and their opinion about a laptop to 7 people, and 4 people were common for both products, then only the 4 people were retained in the combined network. As products are combined, the non-common ties are lost, and to ensure equivalence in the process, the number of products per category were kept the same. This resulted in six aggregated data matrices, one for each product type and valence.

Table 3 contains the summary of the estimated network parameters. The data is in line with previously studied social networks, which provides an external validation of the results. The clustering coefficients of all WOM networks are similar to those of other popular social networks like Facebook, Orkut, and Flickr, typically ranging between .200 and .500 (Hardiman and Katzir, 2013; Ugander et al., 2011). The WOM densities are similar to the densities of the IMDB

social network (18%), and peer-to-peer file sharing (25%) (Melancon, 2006). The average degrees of the WOM networks are similar to the degrees of YouTube (8.5), email network (9.6), and services network (13.5) (Libai et. al, 2013). Finally, the small-world test ratio, S , if greater than one, indicates that a network is a small-world (Humphries and Gurney, 2008). The S ratio for the social network is 1.49, which is similar to the ratio of 1.34 of student social networks from previous studies (Humphries and Gurney, 2008). Regarding the social network, averaging the two ties strength questions revealed that 82% of the ties had strength of less than two, indicating no relationship; 1.3% of the ties were larger than 6. indicating the strongest relationship; and the reaming 17% of the ties had moderate values.

Table 3: WOM Networks Estimates

	Positive WOM					Negative WOM				
	L	C	Deg.	Dens. (%)	S^a	L	C	Deg.	Dens. (%)	S^a
Search Products										
Windbreaker Jacket	2.30	0.35	8.34	13.90	2.08	2.43	0.29	7.42	12.40	1.96
Laptop Computer	2.13	0.36	11.18	18.60	1.55	1.98	0.45	13.38	22.30	1.61
Digital Camera	2.46	0.24	7.36	12.30	1.63	2.04	0.45	12.62	21.00	1.70
<i>Combined Search Products</i>	<i>3.22</i>	<i>0.22</i>	<i>4.32</i>	<i>7.20</i>	<i>2.68</i>	<i>2.79</i>	<i>0.24</i>	<i>5.15</i>	<i>8.60</i>	<i>2.45</i>
Experience Products										
Restaurant Meal	2.03	0.48	13.49	22.50	1.65	2.15	0.37	11.15	18.60	1.55
Vacation	2.14	0.41	11.07	18.40	1.75	2.22	0.26	8.72	14.50	1.49
Movie	1.94	0.49	14.74	25.60	1.47	2.16	0.34	10.74	17.90	1.53
<i>Combined Experience Products</i>	<i>2.18</i>	<i>0.43</i>	<i>8.74</i>	<i>17.90</i>	<i>1.88</i>	<i>2.78</i>	<i>0.26</i>	<i>6.23</i>	<i>10.40</i>	<i>2.03</i>
Credence Products										
Tax Advice/Preparation	2.87	0.29	5.44	9.10	2.63	2.98	0.27	4.95	8.30	2.76
Home Maintenance	2.59	0.30	6.41	10.70	2.36	2.75	0.25	6.16	10.30	1.97
Auto Repair	2.65	0.19	6.51	10.80	1.44	2.74	0.17	5.71	9.50	1.54
<i>Combined Credence Products</i>	<i>3.56</i>	<i>0.16</i>	<i>2.84</i>	<i>5.20</i>	<i>3.14</i>	<i>4.24</i>	<i>0.13</i>	<i>2.62</i>	<i>5.40</i>	<i>2.00</i>

Social Network Parameters	L	C	Deg.	Dens. (%)	S^a
Social Network (All Ties) ^b	1.90	0.41	13.48	22.50	1.49

NOTES: L – Average path length; C – Clustering Coefficient; Deg. – Degree; Dens. – Density; S – Small-World Test Ratio. The estimates were done through the software Gephi.

^a $S = (C/C_{\text{random}})/(L/L_{\text{random}})$, (Humphries and Gurney, 2008), where $L_{\text{random}} \approx \log(n)/\log(np)$, $C_{\text{random}} \approx p$ - network density, n = network size (Newman, 2003b).

^b Tie strength was measured by two scales coded 1 to 7. Any strength from 2 to 7 was coded as a tie, and 1 indicated a lack of connection.

Research questions 1 and 2 were examined together through the average degree (i.e., to how many people a person shares WOM). WOM was the dependent variable in a general linear mixed model, where product type and product valence were the fixed factors, and each respondent was entered as a random variable. The results are presented in Table 4. The effect sizes represented by η^2 are the highest for product type ($\eta^2 = .43$), followed by respondents ($\eta^2 = .38$) and valence ($\eta^2 = .02$). Compared to WOM about credence products, on average, people are more likely to share WOM about experience products to 4.75 more people, and WOM about search products to 2.01 more people. For WOM valence, on average, people are more likely to share positive WOM to 0.63 more people than negative WOM. The main conclusion is that, after the effect of respondents is removed, product type has a substantially stronger effect on WOM than the valence of the opinion. Note that the relative importance of product types is similar to the results obtained from the preliminary study, which is based on a different sample and methodology. This confirmation provides additional validity for the network analysis approach.

Table 4: General Linear Mixed Model for WOM

Independent Variables		WOM ($R^2=.50$)	Partial Eta Squared, η^2
Intercept		1.26	
Product type	Search	2.01 **	.43
	Experience	4.75 **	
	Credence	0.00 ^a	
Valence	Positive	0.63 *	.02
	Negative	0.00 ^a	
Respondents ^b		-	.38

Notes: ^a Parameter set to zero; ^b Respondents were entered as random variables and the table shows only their cumulative effect size

** Significant at $p < 0.0001$; * Significant at $p < 0.05$

The third research question asks how the WOM networks compare to the underlying social network. To answer that question, a series of paired samples t tests compared the average degree of the social network to the average degrees of all WOM networks, individual or combined. The results revealed that the majority of the products do not activate the entire social network (Table 3), meaning that people talk, on average, to less people than their average social connections. The only exception is sharing a positive opinion about a movie that statistically exceeds the social network, where people are willing to share their opinion, on average, to 14.74 people (i.e., degree), although being connected, on average, to 13.48 people. From the combined WOM networks for search experience and credence products, positive WOM for experience products is the closest to the social network, with people sharing, on average, an opinion to 8.74 other people, but still significantly lower.

The values of the network parameters not specifically included in the research questions also reveal interesting patterns. For example, for the combined networks, on average, people share WOM to 3.4 times more people for positive experience products than for positive credence products (17.90 degrees vs. 5.20 degrees, respectively). At the same time, the clustering coefficients of these two networks (0.43 vs. 0.16) show that people talk less about credence products, even within their social clusters. In a social cluster, on average, it is 2.7 times more likely to hear an opinion about a positive experience product than an opinion about a positive credence product. However, the increased S ratio from 1.88 for positive experience products to 3.14 for positive credence products indicates that talking to people outside their social cluster diminishes much faster than talking within the cluster, which confirms the sensitive nature of WOM for credence products.

Network parameters show a substantial variability, even within the same product type. For example, people share a negative opinion about a digital camera, on average, to 7.36 people, and a negative opinion about a laptop to about 13.38 people. Even for the same products, people can share opinions with different valence quite differently: e.g., positive opinion about a digital camera shared, on average, to 7.36 people, but a negative opinion about a digital camera shared to 12.62 people; and for a windbreaker jacket, the direction is reversed. Overall, Table 3 demonstrates the richness of results when WOM is measured as a network. Although the goal of this study is not to make all possible comparisons, it is evident that a network analysis of WOM can be more informative, which opens the door for new enquiries and research questions.

6. Discussion & Managerial Implications

The network analysis of traditional WOM revealed several results that are useful to practitioners. WOM depends primarily on product type. People are most likely to share opinions about experience products, followed by search products, and credence products. For WOM valence, people are marginally more likely to share positive WOM than negative WOM. Generally, people do not utilize their entire social network for WOM, with the notable exception about positive movie experiences, where WOM can surpass even the social network.

From a practical point of view, the above suggests that WOM management should focus on product type. This does not imply ignoring negative WOM, because negative information weighs more than positive information in the mind of a person (Baumeister et al., 2001). However, managers could utilize product type in WOM by including search, experience, or credence qualities in the message they want people to talk about. As Levitt (1980) famously said,

anything can be differentiated. Going one step further, one can argue that every product has search, experience, or credence qualities to some extent; to increase WOM, managers should focus on providing positive experiences, and to limit WOM, managers should focus on credence qualities. For example, motor oil is a typical credence product, which people naturally do not discuss much. To stimulate a discussion about it, WOM about motor oil should be bound to positive experiences like events, creative commercials, social events, rewards, etc. In this regard, one famous example was the partnership between DuPont and NASCAR, where Jeff Gordon promoted Tyvek on his car. Although for most people, house insulation has credence qualities, people talked about Tyvek due to the positive experience they had during the race. Similarly, if managers want to limit the effect of negative WOM, the failure of the product could be attributed to credence qualities that will be less likely to disseminate. This line of reasoning also can explain why some brands naturally relied on WOM to grow. In the past, brands like Avon and Oriflame, which sell cosmetic products in informal home settings, relied heavily on WOM within a personal network. After participating in home selling events, people could experience the benefits of the products and then follow the social vine to share their impressions. Customers are more prone to share positive WOM because of their other-orientation – concern for other customers and helping them to make an informed purchase decision (Bi, Zhang, and Ha, 2018)

Another practical application of WOM as networks is the rich information extracted from them, which can help companies understand and better manage the WOM traffic of their products. For example, clustered conversations with long path lengths may be an indication that customers are hesitant because there is something they do not understand. They may seem to talk to friends but would not share opinions with less known people. In such cases, the focus should be on clarifying the benefits of the product and motivating conversations with acquaintances and people from other social circles in order for the message to travel farther. One example is digital camera brands, where loyal customers can be designated as ambassadors for promoting the brand on social media or in their everyday lives. A monetary or symbolic compensation (e.g., early access to new products) is offered in return. A WOM with low clustering would indicate an absence of local chatter, which is not an optimal seeding campaign (Garber et al., 2004). Fixing such a problem would require targeting strong social connections, which, for example, could be achieved through referral incentives or group discounts.

The last managerial recommendation is related to data collection. Although we employed the random sampling technique due to the comparison of different products types, the “forest fire” sampling technique is more accurate and easier to implement in practice. Forest fire is similar to a snowball sampling, where a respondent passes surveys to the people they recommended the product to. The “fire” can be started in a social environment natural for the product of interest such as a neighborhood, organization, or a club. Because no connections should be lost, it is recommended to include incentives for taking the survey. Alternatively, random sampling could be applied to smaller networks such as people within organizations or professional groups (e.g., physicians, lawyers), where an approach similar to ours can be used.

7. Limitations and Future Research

Despite the insightful results, the study has limitations and leaves questions for future research. Combining different WOM networks into a single network leads to connections loss, because people naturally talk to different people about different products. We combined three products per category, which resulted in a loss of about 20% or more of the connections. All network parameters of combined products have lower WOM estimates than those of the individual products WOM networks. Although a combination allows generalizations, it still is not clear what the best practice is to deal with the loss of information. More research is needed in that direction to understand the reasonable limit of generalizations. It could be possible to represent a combined result with a corresponding level of uncertainty that reflects the combined deviation of the included networks.

The use of a small sample of university students is a limitation and future studies need to address this gap. It is important to replicate the results in non-student and student samples to see whether the results will change. It also is important to replicate the results in samples of different sized because in network analysis sample size can affect network parameters. For example, the S ratio, the average path length, and the clustering coefficient all increase as a function of network size. To mitigate these effects, the study kept the network segment and the products number constant. Therefore, the network results in Table 3, although very close to previously measured social networks,

have value mainly as a comparative tool for finding differences and they answered the research questions. However, because this is a new approach for studying traditional WOM, replication in different settings are needed.

Several other research questions emerge from this study. First, although the provided managerial recommendations are consistent with the network theory, future research should test their causal validity in real life. For example, how exactly will the WOM network change if the efforts focus on increasing the clustering or on shortening the average path length? To what extent can negative WOM be suppressed if the company message focuses on credence vs. positive experience qualities? Second, product type is found to be the strongest predictor of WOM, but are there other characteristics that can be more predictive? For example, WOM networks might be affected more by industry, geographic region, or culture. Finally, it is worth exploring if positive and negative opinions tend to form positive and negative silos in the society. It is known that on social media, people self-separate in networks based on their political orientations. Do WOM networks exhibit such tendencies?

Overall, there are many questions the network approach can answer, and many more to be researched. We hope its introduction to tradition WOM will inspire more interest. Considering the strong standing of traditional WOM and the challenges in assessing the effectiveness of WOM campaigns in general (Libai et al., 2013), network analysis can be a new and helpful tool to improve the management of traditional WOM campaigns.

References

- Alexandrov, A., Lilly, B., and Babakus, E. 2013. The effects of social-and self-motives on the intentions to share positive and negative word of mouth. *Journal of the Academy of Marketing Science*, 41 (5), pp 531–546.
- Anderson, B. S., Butts, C., and Carley, K. 1999. The interaction of size and density with graph-level indices. *Social Networks*, 21, pp 239–267.
- Anderson, E. 1998. Customer satisfaction and word of mouth. *Journal of Service Research*, 1, pp 5–17.
- Anderson, E. and Simester, D. 2014. Reviews without a purchase: Low ratings, loyal customers, and deception. *Journal of Marketing Research*, 51, pp 249–269.
- Aral, S. 2014. The problem with online ratings. *MIT Sloan Management Review*, 55, pp 47–52.
- Baker, A., Donthu, N., and Kumar, V. 2016. Investigating how word-of-mouth conversations about brands influence purchase and retransmission intentions. *Journal of Marketing Research*, 53 (2), pp 225–239.
- Baumeister, R., Bratslavsky, E., Finkenauer, C., and Vohs, K. 2001. Bad is stronger than good. *Review of General Psychology*, 5, pp 323–370.
- Berger, J. and Milkman, K. L. 2012. What makes online content viral? *Journal of Marketing Research*, 49 (2), pp 192–205.
- Berger, J. and Iyengar, R. 2013. Communication channels and word of mouth: How the medium shapes the message. *Journal of Consumer Research*, 40 (3), pp 567–79.
- Berger, J. 2014. Word of mouth and interpersonal communication: A review and directions for future research, *Journal of Consumer Psychology*, 24 (4), pp 586–607.
- Berger, J. 2016, *Contagious: Why things catch on*. Simon and Schuster.
- Bi, N.C., Zhang, R. and Ha, L. 2019, Does valence of product review matter? The mediating role of self-effect and third-person effect in sharing YouTube word-of-mouth (vWOM), *Journal of Research in Interactive Marketing*, 13(1), pp. 79–95.
- Brown, J. J., and Reingen, P. H. 1987. Social ties and word-of-mouth referral behavior. *Journal of Consumer Research*, 14(3), pp.350–362.
- Carl, W. 2006. What's all the buzz about? Everyday communication and the relational basis of word-of-mouth and buzz marketing practices. *Management Communication Quarterly*, 19 (4), pp 601–634.
- Carrington, P., Scott, S., and Wasserman, S. 2005. *Models and Methods in Social Network Analysis*. Cambridge, England: Cambridge University Press.
- Cohen, R., Erez, K., Havlin, S., Newman, M., Barabási, A., and Watts, D. J. 2011. Resilience of the internet to random breakdowns, in *The Structure and Dynamics of Networks*, Princeton University Press, pp 507–509.
- Coldewey, D. 2021. Amazon deflects responsibility on fake reviews but admits 200M were blocked last year. TechCrunch. Available at <<https://social.techcrunch.com/2021/06/16/amazon-deflects-responsibility-on-fake-reviews-but-admits-200m-were-blocked-last-year>> [accessed Nov 4, 2021].
- Collinson, P. 2018. Facebook fake review factories uncovered by Which? Investigation. The Guardian. Available at <<https://www.theguardian.com/money/2018/oct/20/facebook-fake-amazon-review-factories-uncovered-which-investigation>> [accessed Nov 4, 2019].
- Confessore, N., Dance, G., Harris, R., and Hansen, M. 2018. The follower factory. The New York Times. < Available at <https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html> > [accessed Nov 4, 2019].

- De Bruyn, A., and Lilien, G. 2008. A multi-stage model of word-of-mouth influence through viral marketing. *International Journal of Research in Marketing*, 25, pp 151–163.
- De Matos, C., and Rossi, C. 2008. Word-of-mouth communications in marketing: A meta-analytic review of the antecedents and moderators. *Journal of the Academy of Marketing Science*, 36, pp 578–596.
- Dwoskin, E. and Timberg, C. 2018. How merchants use Facebook to flood Amazon with fake reviews. *Washington Post*. Available at < https://www.washingtonpost.com/business/economy/how-merchants-secretly-use-facebook-to-flood-amazon-with-fake-reviews/2018/04/23/5dad1e30-4392-11e8-8569-26fda6b404c7_story.html > [accessed Nov 4, 2019].
- East, R., Hammond, K., and Wright, M. 2007. The relative incidence of positive and negative word of mouth: A multi-category study. *International Journal of Research in Marketing*, 24, pp 175–184.
- Eisingerich, A., HaeEun H., Liu, Y., Jia, H., and Bell, S. 2015. Why recommend a brand face-to-face but not on Facebook? How word-of-mouth on online social sites differs from traditional word-of-mouth. *Journal of Consumer Psychology*, 25 (1), pp 120–128.
- Engagement Labs. 2021. Word of mouth remains a juggernaut: A 15-year WOM measurement journey. Available at < <https://blog.engagementlabs.com/word-of-mouth-remains-a-juggernaut-a-15-year-wom-measurement-journey> > [accessed Nov 4, 2021].
- Fay, B. and Larkin, R. 2017. Why online word-of-mouth measures cannot predict brand outcomes offline: Volume, sentiment, sharing, and influence metrics yield scant online–offline WOM correlations. *Journal of Advertising Research*, 57 (2), pp 132–143.
- Ford, G., Smith, D., and Swasy, J.L. 1988. An empirical test of the search, experience and credence attributes framework. *Advances in Consumer Research*, 15 (1), pp 239–243.
- Frank, O. 2002. Using centrality modeling in network surveys. *Social Networks*, 24, pp 385–394.
- Frenzen, J. and Nakamoto, K. 1993. Structure, cooperation, and flow of market Information. *Journal of Consumer Research*, 20, pp 360–375.
- Fulgoni, G. M. and Lipsman, A. 2015, “Digital word of mouth and its offline amplification: A holistic approach to leveraging and amplifying all forms of WOM,” *Journal of advertising research*, 55 (1), 18–21.
- Galaskiewicz, J. 1991. Estimating point centrality using different network sampling techniques. *Social Networks*, 13, pp 347–386.
- Garber, T., Goldenberg, J., Libai, B., and Muller, E. 2004. From density to destiny: Using spatial dimension of sales data for early prediction of new product success. *Marketing Science*, 23, pp 419–428.
- Girard, T. and Dion, P. 2010. Validating the search, experience, and credence product classification framework, *Journal of Business Research*, 63 (9–10), pp 1079–1087.
- Hanna, N., Wozniak, R., and Hanna, M. 2013. *Consumer Behavior: An Applied Approach*. Dubuque, IA: Kendall Hunt Publishing Company.
- Hardiman, S., and Katzir, L. 2013. Estimating clustering coefficients and size of social networks via random walk. In *Proceedings of the 22nd International Conference on World Wide Web*, International World Wide Web Conferences Steering Committee, pp 539–550.
- Hinz, O., Schulze, C., and Takac, C. 2014. New product adoption in social networks: Why direction matters, *Journal of Business Research*, 67 (1), pp 2836–2844.
- Hinz, O., Skiera, B., Barrot, C., and Becker, J. 2011. Seeding strategies for viral marketing: An empirical comparison. *Journal of Marketing*, 75 (6), pp 55–71.
- Humphries, M., and Gurney, K. 2008. Network ‘small-world-ness’: A quantitative method for determining canonical network equivalence. *PLoS One*, 3(4), e0002051.
- Iyengar, R., Van den Bulte, C., and Valente, T. 2011. Opinion leadership and social contagion in new product diffusion. *Marketing Science*, 30, pp 195–212.
- Jansen, B., Zhang, M. Sobel, K. and Chowdury, A. 2009. Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, 60, pp 2169–2188.
- Katz, E. and Lazarsfeld, P. 1955, *Personal influence*, Free Pr.
- Kossinets, G. 2006. Effects of missing data in social networks. *Social Networks*, 28, pp 247–268.
- Levitt, T. 1980. *Marketing Success through Differentiation of anything*. Cambridge, MA: Harvard University Graduate School of Business Administration.
- Leskovec, J. and Faloutsos, C. 2006. Sampling from large graphs. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp 631–636.
- Li, Feng and Timon C. Du 2017. The effectiveness of word of mouth in offline and online social networks. *Expert Systems with Applications*, 88, pp 338–51.
- Libai, B., Muller, E. and Peres, R. 2013. Decomposing the value of word-of-mouth seeding programs: Acceleration versus expansion. *Journal of Marketing Research*, 50, pp 161–176.
- Litvin, S.W., Goldsmith, R.E. and Pan, B. 2008, Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, 29(3), pp 458–468.

- Liu, H., Jayawardhena, C., Osburg, V.-S., Yoganathan, V., & Cartwright, S. 2021. Social sharing of consumption emotion in electronic word of mouth (eWOM): A cross-media perspective. *Journal of Business Research*, 132, pp 208–220.
- Liu-Thompkins, Y. 2012. Seeding viral content: The role of message and network factors. *Journal of Advertising Research*, 52, pp 465–478.
- Luca, M., and Zervas, G. 2016. Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science*, 62, pp 3412–3427.
- Martin, W. C. and Lueg, J. E. 2013. Modeling word-of-mouth usage. *Journal of Business Research*, 66, pp 801–808.
- Mayzlin, D., Dover, Y., and Chevalier, J. 2014. Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, 104, pp 2421–2455.
- McCarty, C., Killworth, P. D., and Rennell, J. 2007. Impact of methods for reducing respondent burden on personal network structural measures. *Social Networks*, 29, pp 300–315.
- Melancon, G. 2006 (May). Just how dense are dense graphs in the real world? A methodological note. In *Proceedings of the 2006 AVI Workshop on Beyond Time and Errors: Novel Evaluation Methods for Information Visualization*, pp 1–7.
- Meuter, M.L., McCabe, D.B. and Curran, J.M., 2013. Electronic word-of-mouth versus interpersonal word-of-mouth: are all forms of word-of-mouth equally influential? *Services Marketing Quarterly*, 34(3), pp 240–256.
- Mitra, K., Reiss, M.C., and Capella, L.M. 1999. An examination of perceived risk, information search and behavioral intentions in search, experience and credence services. *The Journal of Services Marketing*, 13, pp 208–228.
- Murray, K.B. 1991. A test of services marketing theory: consumer information acquisition activities. *Journal of Marketing*, 55, pp 10–25.
- Musil, S. 2021. Amazon asks social media companies to help stem spread of fake reviews, *CNET*. Available at < <https://www.cnet.com/news/amazon-asks-social-media-companies-to-help-stem-spread-of-fake-reviews> > [accessed Nov 4, 2021].
- Newman, M.E. 2003a. The structure and function of complex networks. *Society for Industrial and Applied Mathematics*, 45, pp 167–256.
- Newman, M.E. 2003b. Random graphs as models of networks. In Bornholdt, S. and Schuster, H.G., *Handbook of Graphs and Networks*, pp 35–65. Weinheim, Germany: Wiley-VCH GmbH & Co. KGaA.
- Newman, M.E. and Park, J. 2003. Why social networks are different from other types of networks. *Physical Review*, E68, 036122.
- Neumann, J. 2017. New study: How buyers research products and where vendors are falling short. *TrustRadius for Vendors*. Available at < <https://vendors.trustradius.com/how-buyers-research-products-and-where-vendors-are-falling-short/> > [accessed February 24, 2018].
- Nguyen, N. 2021. Fake reviews and inflated ratings are still a problem for Amazon. *Wall Street Journal*. Available at < <https://www.wsj.com/articles/fake-reviews-and-inflated-ratings-are-still-a-problem-for-amazon-11623587313> > [accessed Nov 4, 2021].
- Nitzan, I. and Libai, B. 2011. Social effects on customer retention. *Journal of Marketing*, 75 (6), pp 24–38.
- Pagani, G. A. and Aiello, M. 2013. The power grid as a complex network: A survey. *Physica A: Statistical Mechanics and its Applications*, 392 (11), 2688–2700.
- Rosario, A., Sotgiu, F., De Valck, K., and Bijmolt, T. 2016. The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *Journal of Marketing Research*, 53 (3), pp 297–318.
- Schmitt, P., Skiera, B., and Van den Bulte, C. 2011. Referral programs and customer value. *Journal of marketing*, 75 (1), pp 46–59.
- Silverman, G. 2001, *Secrets of word-of-mouth marketing*. New York: AMACOM Books.
- Smith, T., Coyle, J., Lightfoot, E., and Scott, A. 2007. Reconsidering models of influence: The relationship between consumer social networks and word-of-mouth effectiveness. *Journal of Advertising Research*, 47, pp 387–397.
- Sterling, G. 2018. Study finds 61 percent of electronics reviews on Amazon are ‘fake,’ Marketing Land. Available at < <https://marketingland.com/study-finds-61-percent-of-electronics-reviews-on-amazon-are-fake-254055> > [accessed Nov 4, 2019].
- Travers, J. and Milgram, S. 1967. The small world problem. *Psychology Today*, 1 (1), pp 61–67.
- Trusov, M., Bucklin, R.E., and Pauwels, K. 2009. Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site. *Journal of Marketing*, 73, pp 90–102.
- Villanueva, J., Yoo, S., and Hanssens, D.M. 2008. The impact of marketing-induced versus word-of-mouth customer acquisition on customer equity growth. *Journal of Marketing Research*, 45, pp 48–59.
- Ugander, J., Karrer, B., Backstrom, L., and Marlow, C. 2011. The anatomy of the Facebook social graph. *arXiv Preprint:1111.4503*.
- Wasserman, S. and Faust, K. 1994. *Social network Analysis: Methods and Applications*. Cambridge, England: Cambridge University Press.
- Watts, D. 1999. Networks, dynamics, and the small-world phenomenon. *The American Journal of Sociology*, 105, pp 493–427.
- Watts, D. 2003. *Six Degrees: The Science of a Connected Age*. New York: W.W. Norton & Company.

- Watts, D. and Dodds, P. 2007. Influentials, networks, and public opinion formation. *Journal of Consumer Research*, 34, pp 441–458.
- Watts, D. and Strogatz, S.H. 1998. Collective dynamics of small-world networks. *Nature*, 393, pp 440–442.
- Zeithaml, V. 1981. How consumer evaluation processes differ between goods and services. In *Marketing Services*, (ed. by J. H. Donnelly and W. R. George), pp 186-189. Chicago: American Marketing Association.
- Zhang, J., Liu, Y., and Chen, Y. 2015. Social learning in networks of friends versus strangers. *Marketing Science*, 34, pp 573–589.
- Van Doorn, J., Lemon, K. N., Mittal, V., Nass, S., Pick, D., Pirner, P., & Verhoef, P. C. 2010. Customer engagement behavior: Theoretical foundations and research directions. *Journal of Service Research*, 13(3), pp.253–266.

Applicability of Grounded Theory Methodology in International Relations Studies

Faies Kh Jafar

Project Management and Civil Infrastructure Systems Group

The University of Trinidad and Tobago

Faies.tt@gmail.com

Abstract: Engaging the Grounded Theory Methodology (GTM) in International Relation (IR) studies can be a challenging choice for researchers and Ph.D. students. Considering that scholars in the IR field are familiar with certain traditional methodologies, the notion of engaging other approaches such as the GTM may attract strong ontological and epistemological questions. This article contends that the constructivist - and to a certain extent the pragmatist - versions of GTM can be a successful research methodology in IR qualitative research. Such choice, however, is constrained by a set of conditions. Guided by existing literature, these conditions were identified and discussed in the context of IR qualitative research norms and then applied in Foreign Policy (FP) decision-making process, a well-known approach in studying FP as a subdiscipline in IR. The article concludes that despite certain limitations, the methodology can be an outstanding option for IR qualitative research.

Keywords: Grounded theory, qualitative research, IR studies, interdisciplinary research

1. Introduction

IR discipline goes beyond the study of formal relations between states. It includes the study of all actions and interactions among state agencies, non-state actors, institutions, business firms, social media, and others. The phenomenon that IR is concerned with is complex because of its association with different social, political, and economic realities including trade negotiations, global and regional governance formation, and wars (Ainley and Brown, 2005). The increased need to analyze international politics and to understand the nature of relations among nations played an important role in the establishment of the IR discipline a century ago (Koliopoulos, 2017, p.1, 3 & 11).

The nature of the IR research made it an academic autonomous discipline. This autonomy did not restrict the discipline from continuously expanding its interactions to include emerging areas of studies such as economy, sociology, history, and political science. Yet, notwithstanding such interactions, IR studies in the last few decades have become more independent by having “*a distinct set of theories, methodology, and subject matter*” (Chatterjee, 2010, p.3). This autonomy is also characterized by its units of analysis (conflicts, drivers of international political economy, interactions with international organizations among others), and the domains/levels of analysis (powerful individuals and groups, state, and international system). IR scholarship also have a variety of subject matters such as conflicts, negotiations, foreign policy actions, and international trade treaties. The growing independency, however, did not prevent methodological traditions such as phenomenology, case study and research-specific approach to continue to be adopted by IR scholars, however, GTM is glaringly absent from these options.

This expansion of the use of GTM in other disciplines motivated the conduct of the current study. As such an aim is set to investigate as how GTM can be a successful methodology in IR qualitative research. This aim is fulfilled by achieving set of objectives. Firstly, IR methodologies are briefly overviewed; and secondly, GTM is succinctly explained and introduced. Thirdly, those conditions of the methodology that converge with the norms of IR are reviewed and discussed in the context of research paradigm, the role of the literature, the construction of epistemic frameworks, and method of data collection and sampling. And before applying the methodology to a specific area in IR i.e., the Foreign Policy (FP) formulation and implementation process, the paper incorporated the level of analysis problem with GTM, a keystone in any IR study. These objectives are conducive for (a) informing the ontological and epistemological assumptions of the study; (b) clarify the bases of methodological choice; and (c) ultimately, argue that GTM is worthy for IR researchers to consider although with some limitations.

2. Methodologies in IR studies: an Overview

The many great debates in IR reflected its theoretical and methodological evolution. Engaging certain classical theories such as realism and the rise of constructivism inspired IR scholars to elevate the debates to different levels. New debates continued to emerge and has reached the current, the so-called fifth level, in which discourses are centered on linearity and nonlinearity paradigms (Kavalski, 2007, p.445). The epistemology of the theoretical debates reflected preferences towards choices of methods. For instance, realists adopted positivistic methods, constructivists were in favor of the adoption of critical approaches, while behavioralists during the 1950s-1960s advocated the application of naturalist science methods (Benneyworth, 2010).

Methodologies in IR can be categorized into traditional, behavioral, positivist, and postpositivist. Traditional methodologies relied conspicuously on philosophy, international political history, and its epistemology was heavily dependent on experience and logic. In some cases, it converged with post-positivism, particularly when adopting critical analysis. The behavioral methodologies which favor the systematic study of IR phenomena, *"wanted verifiable hypotheses and quantifiable data,"* the aim of which is to produce an explanatory rather than normative research. The positivists who were motivated by the powerful assumption of the existence of objective reality, went a further step by adopting empirical approaches to construct a testable hypothesis. The 'objectivity' assumption was challenged by the emergence of new methodologies which claimed the role of subjectivity in gathering and analyzing data and interpreting studies' results. The latter highlights the birth of post-positivism which included many approaches such as constructivism and postmodernism. Post-positivism reestablished attention to the critical and the normative approaches in IR (Chatterjee, 2010, p.4).

Within postpositivist qualitative research, IR borrows different approaches from social science although when surveying the existing literature, it appears that the IR discipline lacks rigorous self-contained and independent approaches. Add to this, *"the relative absence of research methods textbook"* (Lamont, 2015, p.14) imposed further constraints on IR scholars to broaden their investigation and, therefore, they continued to be restricted by methodological borders that create challenges to engage further theoretical debates.

Among the frequently adopted approaches, particularly in qualitative IR research, are phenomenology when the formation and transmission of certain norms are studied, case studies when the data collection and the analysis are centered on a specific case such as causes of civil conflicts (Creswell, 2011), and discourse and content analysis, but not grounded theory tradition!

Grounded Theory Methodology (GTM) appears to be well established in different studies (Bryant and Charmaz, 2007a, p.1), but not in IR. Since 1967, debates involving the methodology were/continued to be centered around the emerging postpositivist paradigms regarding ontological and epistemological assumptions, and the associated methods of analysis. At first glance, these debates may, normatively speaking, give no reason to exclude IR qualitative studies from adopting GTM. However, the reality on the ground suggests a clear absenteeism of the methodology in IR research. This explicit void in knowledge necessitates answering many questions, the first being the obvious question of 'why' which, due to lack of relevant studies, is highly normative while other questions are the Socratic type such as how the methodology can be engaged in IR, on what basis, and under what conditions, and what kind of challenges and limitation such choice may imply. There is available literature, although not particular to IR, that can be engaged for comparative and critique purposes for these questions except for the question of why. To avoid introducing a highly normative paper, the article does not attempt to answer the latter question.

3. The Grounded Theory Methodology: An Overview

Grounded theory is a methodology commonly used in social science. The main objective of the methodology is to extract and construct a theory grounded in the data. The central premise of the methodology is that a theory about an epistemic problem/gap is institutionalized in the categorical data. By following certain systematic guidelines in gathering, synthesizing, and analyzing these data set of concepts that are logically bonded, a theory emerges and fills this gap. The methodology which adopts the inductive reasoning strategy, in addition, found its applications in a variety of fields of studies such as psychology and health care studies.

The emergence of the methodology returns to 1967 when authors Barney Glaser and Anselm Strauss published their book titled 'The Discovery of Grounded Theory'. The two scholars attempted to respond to ongoing criticism that qualitative research lacked positivist approaches with a particular reference to data collection,

data validation, and methods of analysis (Glaser and Strauss, 1967, p.17). Since then, the two authors' 'invention' of the GTM became another tradition in qualitative studies (Evelyn, 1987). Following their publication, this methodology witnessed the emergence of two other versions: the post-positivist and the constructivist (Levers, 2013). These developments reflected the different ontological and epistemological premises among scholars apart from their aim of flexing the monolithic classical, the so-called traditional version. The emergence of new versions allowed the researchers to expand, adopt and engage the GTM in areas of different studies and disciplines.

Despite differences among GTM versions, the methodology looks at the data, breaks it down into parts, conceptualizes it, and then generates higher-level abstracted categories. Constant comparison between concepts ensures that all relevant data are analyzed. The emerging structure of the concepts and categories becomes a framework that conceptualizes the patterns of behavior of the phenomena under investigation. This replaces the conventional approach of constructing a conceptual or theoretical framework prior to data collection. GTM's developers are aimed at designing "*a well-formed bona fide methodology, not an excuse for not having one*" and a methodology that "*is rigorous and tightly procedural from start to finish*" (Glasser, 2014) to capture the most relevant related concepts explaining the investigated phenomenon.

The emphases of the GTM proponents on nursing and health care disciplines, which can be observed in their published books and papers particularly in the 70's and the 80's, did not form ontological, epistemological, and methodological obstacles for researchers to adopt in other areas of studies (Pettigrew, 2000). The methodology emerged as a feasible choice and gained tremendous appeal as a research strategy for many disciplines (Wells, 1995) and in different areas of investigations including tourism, education, management, business, decision-making studies among many others. Exploring the existing literature via advanced search engines, however, revealed that GTM became a choice for many students and scholars. Professors Antony Bryant and Kathy Charmaz underlined the expansion and the commonness of the use of the methodology among doctoral students. "*GTM is currently the most widely used and popular qualitative research method across a wide range of disciplines and subject areas....doctoral students have successfully completed their degree using GTM*" (Bryant and Charmaz, 2007a, p.1).

Parallel to this expansion, intra GTM debates among the different versions continued. These debates are centered around data analysis in terms of structuring the coding process and constructing categories introducing new types of codes, possible different explanations resulted from adopting certain epistemic positions to the types of reasonings such as inductive, deductive, and abductive, and differences in ontological assumptions. The different versions of GTM, however, still share many common grounds such as the methods of sampling, saturation, the objective of extraction of a particular process, the role and (partially) position of the literature review, and the need to explore the nature of interactions among actors engaged in the process (Timonen, Foley and Conlon, 2018, p.3-4). Devoid of engaging with these intra-debates, the article includes differences and commonalities among the GTM versions with respect to their applications and compatibility in IR research. The aim is to broaden the investigation and thus strengthen the paper's argument.

4. GTM-IR and the Research Paradigm: Limitations

Like other disciplines, the research paradigm in IR is critical in directing the investigation. The paradigm influences the study's epistemology, ontology, and method and methodology, and thus, it is vital to articulate the study's paradigm (Gringeri, Barusch and Cambron, 2013, p. 762). In quantitative research, for instance, positivism is associated with scientific methods while qualitative inquiries are associated with constructivism, feminism, pragmatism among other paradigms. On the methodology level, phenomenology, narrative research, ethnography, case studies, and grounded theory offer well-established platforms for the investigation of qualitative research (Creswell and Creswell, 2018, p.259). GTM, and despite being defined as a methodology, it includes a set of methods in cases of sampling, analysis, and trustworthiness techniques. These methods are among the common grounds of the different versions of the GTM. These versions, however, advocate a different paradigm within the qualitative inquiry.

IR qualitative research is pluralistic. It allows the engagement of diverse analyses and approaches in the investigation process. The constant emergence of new levels of inter and intra debates among IR scholars, gives some degree of freedom of methodological choices. These debates, in addition, feature the differences

regarding ontological and epistemological stands. For instance, realists (adopting rational choice approaches and assume objective reality) verses constructivists (critical method and allow subjectivity) debates.

Barney Glaser and Anselm Strauss were enthusiastic to introduce a qualitative research methodology that fits the positivist paradigm (Bryant and Charmaz, 2007b, p.33). Their Classical Grounded Theory (CGT) methodology design is driven by powerful assumptions such as detaching the researcher from the collected data which influence the process of investigation; prior knowledge is not desired/required as it may contaminate the researcher's thinking during the investigation process. Therefore, as the two scholars contend, and despite being qualitative, such an approach helps the researcher to maintain objectivity. The scholars, in addition, explained that collecting qualitative data and the development of categories can follow similar methods to those in quantitative studies, such as surveys, and that qualitative data do indeed have dimensions and properties like numerical data (Glaser and Strauss, 1967, p.193).

Positivists' paradigm is associated with the scientific method, testing hypothesis, measuring, numerical modeling building, and include statistical tests. Therefore, it is highly challenging for CGT enthusiasts to convince other scholars on how sharing certain similarities qualifies their version to become a qualitative and yet positivist at the same time. For IR this is a great challenge as the hegemony of adopting scientific (rational) methodologies in IR in many parts of the world continue to dominate scholars' choices, and this is not the reality of the western school only, but also the case of other emerging schools such as the Chinese (Eun, 2016, p.2-6). On the ontology level and despite the emphases of Barney Glaser and Anselm Strauss on the existence of the independent reality, apprehending it is never perfect. Thus, their version inevitably rests on a post-positivist ontology (Hallberg, 2006, p.7). Add to this, methods associated with positivism are mostly deductive (GTM is inductive) and are in favor of quantitative analysis. This represents a clear pitfall of GTM as a choice by scholars, including IR's. These metatheoretical debate no doubt complicates any IR research proposal as it generates numerous questions by examiners, peer reviewers of papers and may force the researcher to engage with broader unnecessary discussions to promote this version as a positivist one.

Years after working with Barney Glaser, Anselm Strauss worked with Juliet Corbin to flex the methodology. The two authors made it explicit that in GTM, as a qualitative approach of inquiry, subjectivity is inevitable. Their post-positivist interpretivism GTM model is based on the ontological assumption whereby the researcher is actively and constantly engaged with the data (Strauss and Corbin, 2015). On the epistemology level, the two scholars allowed prior knowledge to strengthen the research, in general and data collection, in particular. Therefore, on the paradigm level, Strauss and Corbin's approach forms no obstacle or variation to those in IR stated previously.

Professor Kathy Charmaz went a further step when she engaged the constructivist perspective, the main pillar of the fourth level debates in IR. Charmaz's version can be the optimum option when the subjective reality is well understood through the concept of 'knowledge is constructed'. This leads to a theory-building rather than accepting that reality exists independently, and thus, the role of the researcher is not to extract it, but to construct it (Charmaz, 2006, p.27). Constructivism brought a unique perspective to international politics' studies beyond what the realists' school offered. It opened a new way for interpretations which created a distinct leap in how IR is approached (Hopf, 1998) and established the path to the fourth debate in IR (Kavalski, 2007, p.449).

The synthesis of this section emerged from ongoing inter-paradigm debates among the different versions of GTM. It suggests some limitations when choosing a particular version of the methodology. Even though these debates are consistent with other disciplines in social science, the case in IR research seems delimited to constructivists' approach/s. Although paradigm choice is philosophical in nature, it will be challenging for an IR research proposal to be convincing when adopting GTM with an assumption of consistency with positivism.

5. The Position and Role of the Literature Review, and Epistemic Frameworks

When Barney Glaser and Anselm Strauss introduced their version of CGT, they emphasized that developing a theory from data would be more successful without a *priori* assumptions (Glaser and Strauss, 1967, p.6 & 29). Although the two authors admitted that researchers do not approach reality as a *tabula rasa* (ibid, 1967, p.3), they discourage researchers from conducting literature reviews at an early stage of the study. Years later, Glaser's pointed out that researchers should not "worry" since "*literature isn't going nowhere*" (Glaser, 2010) because it will emerge at a later stage of the study. For IR research, this is a problem since the norm, regardless

of the study paradigm or methods (qualitative, quantitative, or mixed), is positioning the literature review, like other social science studies, at an early part of the thesis. In addition, being an interdisciplinary field, IR literature may introduce and open an early debate as to how to scan and filter the most relevant prior scholarships of the many relevant subfields that can elucidate and identify the gap in knowledge and help to set the study questions. This problem may impose constraints on Ph.D. and scholars in attempting to convince examiners and proposal funding agencies (Dunne, 2011).

Charmaz adopted an opposite position to CGT based on the impossible separation of existing and previous knowledge of the researcher and the investigation process. She maintained that researchers should not let this knowledge have control over the investigation process and to allow the free emergence of themes and categories (Sebastian, 2019, p.4). Constructivist GTM, in other words, allows the establishment of epistemic frameworks and by extension, the conduct of literature review in advance which is consistent with IR research norms. Charmaz's approach, in addition, encourages reflexivity, which can enrich the study with insights and broadening inter and intra debates among IR theories. Studies show that in the last few decades, reflexivity in IR scholarly has risen in prominence (Hamati-Ataya, 2020). The antithesis to this view, however, is that reflexivity will encourage subjectivity, a point that obviously brings the objectivity of explanations of the study into questions.

The literature review, in addition, plays an important role in articulating the research's questions (Creswell and Creswell, 2018, p.162). Therefore, delaying the review of existing empirical studies, as CGT emphasized and to some extent, the pragmatist version of what Corbin and Strauss suggested, may become problematic. The latter's version adopted a vague position on the issue of development of the research question at an early stage of the study. The objective is to let the emergence of data have its role in generating the most relevant questions. In the constructivist GTM, Charmaz explained that data gathering is influenced by the questions, for instance, by the type of data and method of sampling (Charmaz, 2006, p.15). Delaying the development of the research questions similar to the literature review may contradict the supervisory committees' expectations and hence, forms challenges to the researcher and raises all the questions indicated above.

On the epistemic frameworks, Barney and Strauss informed that since the aim is to develop a theory, having a theoretical framework is not part of their methodology. Initial data analysis guides the researcher to make early judgments without the need for frameworks. The authors maintained that researchers could begin with a *"partial framework"* of a set of concepts with the aim of *"designating a few principals or gross features of the structure and processes in the situations that he will study"* (Glaser and Strauss, 1967, p.45). Such a position may demotivate IR scholars as conceptual and theoretical frameworks have a distinct place in their research because of their enabling power of analyzing different issues, organizing thoughts, and setting the most relevant research questions (see for example McClellan 2015; Kavalski, 2007; Klotz, 2008; Curini and Franzese, 2020; Bennion, 2015). The conceptual framework in qualitative research represents the vertebra of the research epistemology and it is subjected to two general conditions. Firstly, it must be convincing as being the most appropriate/relevant to explain and interpret the analyzed data. Secondly, it links the framework to the ontological assumptions and the research methodology in a compatible way to integrate the study's paradigm.

As a pluralistic field, IR research is rich with theories and depends on many concepts from different disciplines. Therefore, the absence of a conceptual framework may raise analytical and methodological questions particularly during the early stage of the research. To prevail over a such challenge, it might be necessary to adopt the spiral relation promoted by Charmaz to open the way among data collection, theoretical sampling, emerging findings, and discussions (the emerging theory) on one hand, and the constructed conceptual framework on the other. The framework will be subjected to constant evolution until the emerging framework reaches the point of equilibrium where relevance to the findings and the discussions becomes highly evidenced (Charmaz, 2006, p.24, 76 & 101), and such a reflexive approach does not contradict emerging IR qualitative research norms (Hamati-Ataya, 2020).

Corbin and Strauss, however, adopted a midground position between CGT and the Constructivist. They established a normative view that allows prior knowledge before and during the data collection to enhance sensitivity, to confirm the significance of initial concepts that appeared in the literature for better comparison with the emerging concepts. On the theoretical frameworks, the authors discouraged their use since the ultimate objective is to develop a theory rather than testing those in existence. For IR scholars, this suggests that it is left to them, with some flexibility, to choose to have theoretical frameworks since Corbin and Strauss do not

have a dogmatic position in this regard. This flexibility is necessary for IR scholars and students to “*meet the requirements and expectations of committee members and universities*” who are in favor of maintaining IR research traditions (Strauss and Corbin, 2015, p. 50- 52 & 314).

6. Data Collection, Sampling, Coding and Analysis, in the Context of IR Research

In GTM primary data, irrespective of the methodology version, relies extensively on interviews. The methodology also allows for other sources of data (secondary) such as participant observations, documentary sources, videos, documents, diaries, memoirs, journal papers among many others (Glaser and Strauss, 1967, p.163 & 175; Strauss and Corbin, 2015, p.37 & 43; Charmaz, 2006, p.13-40). These sources of data are consistent and compatible with IR qualitative research, for instance when conducting elite interviews and analyzing official documents in subdisciplines such as security studies and foreign policy (Leander, 2008, p.5 & 24). Despite such compatibility, the complex relationship between the data collection and the research questions may suggest a clear disadvantage of the CGT. In addition, the level of analysis, which is a very well-known pillar in any IR research may add other constraints. This issue is discussed in a subsequent section.

GTM promotes theoretical sampling. It is the process of choosing the most relevant participants to gather data. The degree of sampling's relevancy is guided by analysis which is conducted along with data gathering (Auerbach and Silverstein, 2003, p.18). Researchers however may start with what they perceive as the most appropriate samples and then proceed to theoretical sampling until saturation. The latter is reached when no new categories emerge from the analysis and thus, collecting additional data becomes redundant. In ‘The SAGE Handbook of Research Methods in Political Science and International Relations’, author Virginie Van Ingelgom, advocated Glaser and Strauss's principle of sampling saturation to IR qualitative research “*to uncover diversity and context, rather than a large number of participants with the same type of experience*” (Ingelgom, 2020, p.1196). Therefore, if the collected data sourced from interviewees in the IR qualitative research, normatively speaking, theoretical sampling approach allows the diverse IR phenomenon to be captured within well-defined boundaries and simplifies the process of borrowing/filtering of many concepts from other disciplines.

Being part of the methodology, for IR research, data collection and sampling form a new understanding because both data collection and sampling are methods, not methodologies. IR's studies, like others, mandate having methods. In any research proposal, having some of these methods as a part of the methodology may raise questions and concerns by committees and Institutional Review Boards (IRBs). The IRB often follows certain standards and protocols to evaluate and approve research proposals (Strauss and Corbin, 2015, p. 42). Therefore, stating and clarifying this issue is essential for successful proposals.

GTM, in addition, consists of coding and analysis methods. Each version adopted a certain coding process. The shared zone among the versions' coding is allowing constant comparison between emerging codes (Sebastian, 2019, p.4) which develops categories signifying patterns that describe the process under investigation. This ‘process’ forms a major part of what IR qualitative research is aiming at, i.e. the extraction of different factors associated with the process. This falls in the heart of IR constructivists' research (Theys, 2018). However, the generalizing judgment of the compatibility of the methodology to IR qualitative research necessitates the answer to an important and leading question related to the levels of analysis.

7. The level of Analysis Challenges and GTM

Methodologically, and in terms of levels of analysis, IR studies are fundamentally different from other areas of social sciences (Singer, 1961; Finkel and Strauss, 2012), particularly at the system level. Choosing the appropriate level increases the “*capacity to explain the relationships among the phenomena under investigation*” (Singer, 1961, p.79). It also helps to determine the possible/particular unit/units of analysis. The IR research mandates an early decision on the level of analysis, the domain in which investigations are carried out. Engaging the level of analysis, in addition, grants researchers some degree of freedom from constantly attempting to conceptualize what international ‘is’ by allowing them to shift the epistemology and the international system from merely an ontological into a methodological (Temby, 2013, p.18).

In his essay titled ‘International Conflict: Three Levels of Analysis’, David Singer (Singer, 1960, p.461) argued that in IR studies and despite the need for the three levels of analysis, it is essential to determine the unit of analysis within the particular level. The individuals' levels, for instance, is not an independent variable but rather those conceptualized processes and events e.g., in the form of foreign policies by decision-makers. These units are to be treated as key variables that explain the way/how the IR phenomenon is perceived by the policymakers. In

this way, Singer shifted the role and assumptions of levels of analysis from being ontological as Kenneth Waltz proposed (Waltz, 1959), into a methodological. Singer's point, however, might be more effective at the individual and state levels. A study with an aim to develop a phenomenological model, for instant, mandates gathering data relevant to policymaker's perceptions, lived experience, personal observations, and belief (Singer, 1961, p.88). Studies with such aims and at these levels, therefore, resonate constructively with GTM features. On the system level, however, acquiring this type of data may not be available or feasible. At this level, common data in qualitative IR research includes documents, audio and video records, international events, interviews already available in the media of high-profile individuals e.g. heads of international organizations, etc. GTM does allow for these data to be included as a secondary source particularly for literature comparison, but the primary data, as stated previously, are gathered and analyzed verbatim from participants' interviews. This also denotes good news for IR research at individuals' level since it represents a shared zone with other fields in social studies engaged successfully GTM since 1967.

Determining/choosing the level of analysis whether at the system, state, or individuals, therefore, plays an important role in defining the type of data to be acquired and analyzed; it is an important factor that directs the study's finding. For instance, categorical data related to the human agency, or the political system might be inadequate to explain the associated phenomena at the system level. Similarly, when the research problem is associated with large-scale events such as inter-state conflicts then the system level appears to be the obvious choice. However, the researchers may look at it differently when the scope of the study changes, for example, if the purpose of the study is to understand the relation between the leadership style and the regional/international conflicts then the level choice becomes problematic and may require a more sophisticated approach to address the multiple level compatibility issues.

When the prime unit of analysis is a phenomenon and a process takes place at/within the state level, the scenario seems less rigid. The collected data, in this case, is well-suited for the coding and the constant comparison which the methodology provides (Strauss and Corbin, 1990). It is 'well-suited' because further shrinkage of the domain of analysis facilitates the extraction of the concepts by capturing its forming components. This is the foundation upon which the grounded theory is constructed.

8. Engaging GTM in IR Subfields: Example of Studying Foreign Policy (FP) Formulation and Implementation Decision-Making Process

Engaging the GTM in an IR study at the individual, groups, and state levels opens the way for upgrading studies from substantive areas into system-level theories. For example, studying feminism and foreign policymaking may explain certain aspects of state behavior that are rooted in its social (domestic) structure, which then finds its way into the international arena in the form of a policy. The gender/feminism issues differ among societies, thus importing concepts based on empirical studies conducted elsewhere may not be helpful to understand the 'state-specific' complex gender-FP relation because of their inherited historical, social, and socio-political distinguishing factors (Bordo, 1990). The latter, in addition, highlights a level of complexity that makes the development of a testable hypothesis almost impossible; therefore, investigating this issue inductively at the individual, groups, and maybe state levels becomes an unavoidable strategy. The emerging theory can then be taken to a higher level by conducting comparative studies between the states.

As a subfield of IR, FP studies are approached differently. The study of the decision-making process of the FP formulation and implementation is one approach whereby the research aims at drawing an holistic picture of the process that drives the state seeking opportunity/responding to the external reality. The policy decision is made by powerful individuals and groups in different institutions. Therefore, the level of analysis is limited to the state, groups, and individuals. The FP process includes "*choosing actions outside the state boundary for achieving goals*" by adopting "*means, goals, values and different instruments which the government uses in making relations with different countries*" (Adnan, 2014, p.657). Choosing the appropriate action follows a decision-making process which is the prime unit of analysis during the FP's life cycle and is highly influenced by surrounding socio-economic and socio-cultural factors (Almaas and Johnson, 2015). Thus, the study of the FP decision-making includes three major aspects compelled to be conceptualized: a process, the role of interacting actors, and the influencing factors.

The FP decision-making re/presented in a model/s explains the policy formulation and implementation process. Models are not theories, but a partial representation of them (Leijonhufvud, 1997, p.193). Concepts and

categories emerging from constant coding represent high-level abstractions, i.e. theoretic coding in CGT, axial and selective coding in the pragmatist approach, and focused coding in the constructivist version. These concepts represent high-level categories that emerged from constant coding and comparisons and they consist of many components e.g. lower-level codes. *"There are no simple concepts. Every single concept has components and is defined by them"* and *"there is no concept with only one component"* (Deleuze and Guattari, 1990, p.15). Around these high-level categories and the other emerging concepts, the models of FP formulation and implementation are developed. Models, as stated, are part of the ultimately emerging grounded theory. The latter includes those factors influencing and shaping the decision models and the policymaking actors.

FP decision-making studies are concerned about factors influencing the policy. Some of these factors are existing realities and labeled as objective determinants (Wanjohi, 2011), others are known as subjective. The latter may include leadership styles, psychological, certain socio-cultural, beliefs, etc. GTM which was developed initially by sociologists, no doubt, resonates positively with such FP influencing factors. The methodology's focus, like other disciplines, is not limited to those influencing factors, but on how these factors interact with the policymaking process. These factors which take the forms of categories, concepts, and other components are extracted from processing (coding) the data directly rather than examining patterns of meaning, the common approach of thematic analysis.

Similarly, the role of actors engaged in the policy process at different levels can be extracted. As a qualitative methodology, GTM additionally guides the researcher to conceptualize policy actors' ontology and epistemology i.e. it captures their understanding of the surrounding reality. Their quality of understanding and perceptions of the objective determinants, for instance, is translated into a particular rational, sometimes emotional level of readings which ultimately influence the policy (Singer, 1961, p.86). GTM goes a further step by allowing the comparison of the behavior of policy actors, groups and individuals, at different times (Glaser and Strauss, 1967, p.149). In this way, it allows psychological, socio-cultural, and even belief systems' influence on the FP process to be investigated and theorized.

Notwithstanding this promising picture, conducting FP decision process analysis is not without its limitation. Confining the data at the individuals and groups level suggests that the study's boundary is the particular state's FP. In other words, it promotes a state-specific case study. A comparative study of a group of states' FP decision-making process forms an obvious challenge for the scholar when collecting data in terms of scale and maybe access. For such a case, developing a formal theory becomes a challenge and therefore, researcher may be compelled to shift the study towards building a substantive theory instead (Glaser and Strauss, 1967, p. 149; Strauss and Corbin, 2015, p.50- 52 & 314). The immediate consequence of such a path is limiting the scope into a substantive area of investigation, extensive reliance on primary field data collected by the researcher, and less (if not) engagement of existing empirical data.

On the level of conceptual/theoretical framework, which was explained previously, the study of FP decision-making imposes additional constraints. The subjective factors allow for a wide range of concepts to generate from coding. Psychological issues, for example, open the way to include numerous motivational factors as sources of biases when the policy is selected. A prior explanation of certain psychological concept from empirical studies in the context of policymaking bias and partiality become inescapable. Therefore, CGT's advocating of having a loose or no epistemic framework in advance becomes problematic, while the other versions, especially the constructivist through its invitation to adopt the spiral relations (moving back and forth) can offer a good solution for such a challenge.

Other aspects of the methodology offer no obstacles for the study of FP formulation and implementation and are consistent with the general principles of any qualitative research. This includes a method of sampling, member checking, and reaching the degree of saturation. (Ingelgom, 2020, p.1196).

In this example, three conditions for adapting GTM are met (a) having process and interactions of factors and actors; (b) the level of analysis is limited to groups and individuals with direct implication on the necessary data gathering method (interviews), and (c) there is constant interaction between the process and the surroundings. Despite some limitations, adopting the methodology is possible with a particular reference to the pragmatists and constructivist versions. Furthermore, under certain conditions and purposes explained above, the study of FP decision-making may consider the development of a substantive theory rather than a formal.

Decision-making studies with GTM (see for example Fenwick, Chaboyer and St John, 2012; Halimi, D'Souza and Sullivan-Mort, 2014; King and Snowden, 2020) uphold the above conclusion. Although using different case studies, the associated published scholarly research maintained those key points indicated with FP example in this section. Fenwick, Chaboyer and St John (2012, p. 53 & 59), for instance, found that GTM helped to explore those aspects influencing the decision-making process when studying self-identifying and conceptualizing self-management. Similarly, by adopting GTM, Halimi, D'Souza and Sullivan-Mort (2014, p.263) explored how cognitive judgements of consumers are influenced by their beliefs about the product. For the individuals to reach a decision, they go through simple and complex judgment processes. Exploring and analyzing the decision-making process in these studies emerged from purposeful theoretical sampling (at individuals' level) encouraged by GTM authors, and most importantly, developed themes (for examples: consultation, provide role view, agent for military; hostility, level of involvement, building boundary etc.). These themes explain a 'process of actions' which is (a) the essence of any decision-making study, and (b) the focus of any study conducted with GTM. These actions are taken by powerful agents which explain why such studies naturally are conducted within the level of individuals and groups of humans.

Dissimilarities between FP decision-making and other decision-making studies, in the context of methodology, are mainly the investigation of domain specific interactions and the unit of analysis. This is an advantage for FP and by extension IR scholarship, because existing studies in other disciplines became an establishment or evidence to a successful choice of the GTM in IR studies under the conditions outlined.

9. Conclusion

This paper is aimed at filling a gap in research methodology: the applicability of the grounded theory methodology in international relations studies. Following the dialectical approach, the aim is fulfilled by engaging existing literature. Supported with an example the paper conclusively showed that under certain conditions, embracing constructivist and pragmatist versions of the methodology in IR research is feasible and promising. More specific, the article introduced the main characteristics of IR qualitative research methods and methodologies, and then conducted a comparison with GTM's key components, whereby discussions were established in the areas of consistency and inconsistency between IR qualitative research methodologies and GTM versions.

On the paradigm level, the adoption of pragmatic approach opens the way for IR scholars to be flexible with respect to objectivity claims by CGT by accepting some level of subjectivity. Such paradigm shifts, may promote the assumption of 'theory construction' which is preferable choice when designing constructivist IR research proposals, a major pillar of the fourth level debates in IR.

Discussions regarding the role and position of the literature review and epistemic frameworks suggest that CGT is not consistent with IR qualitative research norms, which is not the case of the other versions. Constructivist GTM, and to some extent the pragmatic embraced a flexible position. For IR research, and to avoid the risk of raising unnecessary questions by IRBs, maintaining the norm of conducting an early literature review and the establishment of conceptual frameworks is possible when employing constructivist and pragmatist versions. Such a choice requires the employment of reflexivity strategies which might form challenges to least experienced researchers.

Methods associated with the three GTM versions such as data collection, sampling, and data coding and analysis form no contradiction to IR research. These methods, in addition, may offer new insights through the data coding process which directs the investigations towards the development of substantive or formal theory. Determining the point of saturation of sampling offered by the methodology is highly important for IR studies due to its nature of being interdisciplinary which allow gathering large amount of data; saturation delimits this problem and hence produce research with more focus scope.

Assessing GTM-IR research compatibility mandates the engagement of the problem of levels of analysis. The main conclusion in this regard is that adopting the methodology is encouraged when the primary data is interviews. This can be the case of individual and to some extent the state levels. At the system level, when the main source of available data are documents, the adoption of any version of the methodology is not recommended. Treating the level of analysis methodologically, David Singer's view is an advantage for GTM

choice since it clarifies and affirms the appropriateness of the selection of the unit/s of analysis and forms of the collected data.

The example of the FP decision-making process illustrates how GTM can be applied in this IR subfield, where conditions of GTM are met. This included the need for extracting an exist process from data provided by policy agents, investigating the nature of their interactions, and possible factors influencing these interactions. The methodology in this case, adds an advantage as it reconfigures the human agency-structure relations. This advantage, however, is not challenge free since breaking down the policy process into finite elements (concepts and their components) opens the way for complicating and broadening the domain of investigation by inviting other disciplines, such as psychology, to become part of the integrated theory.

References

- Adnan, M., 2014. Foreign policy and domestic constraints: a conceptual account. *South Asian Studies*, 29(2), pp.657-675.
- Ainley, K., and Brown, C., 2005. *Understanding international relations*. 3rd ed. New York: Palgrave Macmillan.
- Almaas, A. and Johnson, K., 2015. *Excerpts about Objective Perception: Seeing Things as They Are*. [online] Available at: < <https://www.diamondapproach.org/almaas> > [Accessed 5 January 2021].
- Auerbach, C., and Silverstein, L., 2003. *Qualitative data: an introduction to coding and analysis*. NYU Press: New York.
- Benneyworth, I., 2010. The 'Great Debates' in international relations theory. [online] Available at: < <https://www.e-ir.info/2011/05/20/the-%E2%80%98great-debates%E2%80%99-in-international-relations-theory/> > [Accessed 8 January 2021].
- Bennion, E., A., Experiential education in political science and international relations. In: J. Ishiyama, W., J., Miller, and E., Simon, eds. 2015. *Handbook on teaching and learning in political science and international relations*. Cheltenham, UK: Edward Elgar Publishing. Ch.30.
- Bordo, S., 1990. Feminism, postmodernism, and gender-skepticism. In: L. Nicholson, ed. 1990. *Feminism/Postmodernism*. London: Routledge. Ch.6.
- Bryant, A. and Charmaz, K., 2007a. Grounded theory research: methods and practices. In: A. Bryant and K. Charmaz, eds. 2007. *The SAGE handbook of grounded theory*. Los Angeles: SAGE. Ch.1.
- Bryant, A. and Charmaz, K., 2007b. Feminism, postmodernism, and gender-skepticism. Grounded Theory in Historical Perspective: an Epistemological Account. In: A. Bryant and K. Charmaz, eds. 2007. *The SAGE handbook of grounded theory*. Los Angeles: SAGE. Ch.2.
- Charmaz, K., 2006. *Constructing grounded theory: a practical guide through qualitative analysis*. London: SAGE.
- Chatterjee, A., 2010. *International relations today: Concepts and applications*. India: Dorling Kindersley.
- Corbin, J. and Strauss, A., 1990. *Basics of qualitative research: grounded theory procedures and techniques*. CA: SAGE Publications.
- Creswell, J., W., 2011. *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. 4th ed. Boston: Pearson Education.
- Creswell, J., W. and Creswell, J., D., 2018. *Research design: qualitative, quantitative, and mixed methods approaches*. 5th ed. London: SAGE Publication Ltd.
- Curini, L., and Franzese, R., 2020. *The SAGE handbook of research methods in political science and international relations*. London: SAGE.
- Deleuze, G. and Guattari, F., 1990. *What is philosophy?* New York: Columbia University Press.
- Dunne, C., 2011. The place of the literature review in grounded theory research. *International Journal of Social Research Methodology*, 14(2), pp.111–124.
- Eun, Y., 2016. To what extent is post-positivism 'practised' in International Relations? Evidence from China and the USA. *International Political Science Review*, [online]. Available at: < <https://journals.sagepub.com/doi/10.1177/0192512116642222> > [Accessed 11 January 2021].
- Evelyn, J., 1987. Qualitative research traditions: a review. *American Educational Research Association*, 57(1), pp.1-50.
- Fenwick, C., Chaboyer, W., and St John, W., 2012. Decision-making processes for the self-management of persistent pain: a grounded theory study. *Contemporary Nurse*, 42(1), pp. 53-66. <http://dx.doi.org/10.5172/conu.2012.42.1.53>
- Finkel, E. and Straus, S., 2012. Macro, meso, and micro research on genocide: Gains, shortcomings, and future areas of inquiry. *Genocide Studies and Prevention: An International Journal*, 1(7), pp.55-67.
- Glaser, B. and Strauss, A., 1967. *The discovery of grounded theory*. London: Aldine Transaction.
- Glaser, B., 2010. *The literature review in grounded theory*. [Online video] Available at: < <https://www.youtube.com/watch?v=7S1kJOk3yHk> > [accessed 25 February 2021].
- Glaser, B., 2014. Choosing Grounded Theory. *Grounded Theory review*. [online]. Available at: < <http://groundedtheoryreview.com/2014/12/19/choosing-grounded-theory/> > [Accessed 12 January 2021].
- Gringeri, C., Barusch, A., and Cambron, C., 2013. Examining foundations of qualitative research: a review of social work dissertations, 2008–2010. *Journal of Social Work Education*, 49(4), pp.760-773.
- Halimi, D'Souza and Sullivan-Mort, 2014. Understanding the role of political relations in consumer decision making: a grounded theory approach. *International Journal of Business and Management*, 9(11), pp. 261-273.

- Hallberg, L., 2006. The “core category” of grounded theory: Making constant comparisons. *International Journal of Qualitative Studies on Health and Well-being*, [online]. Available at: < <https://doi.org/10.1080/17482620600858399> > [Accessed 11 January 2021].
- Hamati-Ataya, I., 2020. *Reflexivity and International Relations*. Oxford University Press.
- Ingelgom, V., V., 2020. Focus groups: from qualitative data generation to analysis. In: L., Curini, and R., Franzese, eds. 2020. *The SAGE handbook of research methods in political and international relations*. London: SAGE. Ch.62.
- Kavalski, E., 2007. The fifth debate and the emergence of complex international relations theory: notes on the application of complexity theory to the study of international life. *Cambridge Review of International Affairs*, 20(3), pp.435-454.
- King, E., L., and Snowden, D., L., 2020. Serving on multiple fronts: a grounded theory model of complex decision-making in military mental health care. 250, pp. 1-10. <https://doi.org/10.1016/j.socscimed.2020.112865>
- Klotz, A., 2008. Introduction. In: A., Klotz and D., Prakash, eds. 2008. *Qualitative methods in international relations: a pluralist guide*. New York: Palgrave Macmillan. Ch.1.
- Koliopoulos, C., 2017. International Relations and the Study of History. *Oxford Research Encyclopedias, International Studies*. [online], pp.1-20. Available at: < <https://doi.org/10.1093/acrefore/9780190846626.013.242> >
- Lamont, C., 2015. *Research methods in international relations*. London: SAGE Publication Ltd.
- Leander, A., 2008. Thinking tools. In: A., Klotz and D., Prakash, eds. 2008. *Qualitative methods in international relations: a pluralist guide*. New York: Palgrave Macmillan. Ch.2.
- Leijonhufvud, A., 1997. Models and theories. *Journal of Economic Methodology*, [online], 4(2), pp.193-198. Available at: < <http://dx.doi.org/10.1080/13501789700000014> >
- Levers, D., 2013. Philosophical Paradigms, Grounded Theory, and Perspectives on Emergence. *Sage Journals*. [online]. Available at: < <https://doi.org/10.1177/2158244013517243> > [Accessed 12 December 2020].
- McClellan, E., F., 2015, Best practices in the American undergraduate political science curriculum. In: J. Ishiyama, W., J., Miller, and E., Simon, eds. 2015. *Handbook on teaching and learning in political science and international relations*. Cheltenham, UK: Edward Elgar Publishing. Ch.1.
- Pettigrew, F., 2000. *Ethnography and Grounded Theory: A Happy Marriage?* [online] Available at: < <http://www.acrwebsite.org/search/view-conference-proceedings.aspx?Id=8400> > [Accessed 5 September 2020].
- Sebastian, K., 2019. Distinguishing between the types of grounded theory: classical, interpretive and constructivist. *Journal for Social Thought*, 3(1), pp.1-9.
- Singer, J., D., 1960, Review: international conflict: three levels of analysis. *World Politics*, 12(3), pp.453-461.
- Singer, J., D., 1961. The level-of-analysis problem in international relations. *World Politics*, 14(1), pp.77-92.
- Strauss, A., and Corbin, J., 2015. *Basics of qualitative research: techniques and procedures for developing grounded theory*. 4th ed. LA: Sage Publications.
- Temby, O., 2013. What are levels of analysis and what do they contribute to international relations theory? *Cambridge Review of International Affairs*. [online], pp.1-23. Available at: < <http://dx.doi.org/10.1080/09557571.2013.831032> >
- Theys, S., 2018. Introducing Constructivism in International Relations Theory. [online] Available at: < <https://www.e-ir.info/2018/02/23/introducing-constructivism-in-international-relations-theory/> > [Accessed 5 January 2021].
- Timonen, V., Foley, G., and Conlon, C., 2018. Challenges when using grounded theory: a pragmatic introduction to doing GT research. *International Journal of Qualitative Methods*, 17, pp.1-10.
- Waltz, K., N., 1959. *Man, the state, and war: a theoretical analysis*. New York: Columbia University Press.
- Wanjohi, A., M., 2011. *State's Foreign Policy: Determinants and Constraints*. [online] Available at: < <http://www.kenpro.org/papers/foreign-policy-determinants-constraints.htm> > [Accessed 5 December 2020].
- Wells, K., 1995. The strategy of grounded theory: possibilities and problems. *Social Work Research*, 19(1), pp.33-37.