International Journal of Information Management 58 (2021) 102310

Contents lists available at ScienceDirect

International Journal of Information Management

journal homepage: www.elsevier.com/locate/ijinfomgt

Technical Note

Fuzzy-set Qualitative Comparative Analysis (fsQCA): Guidelines for research practice in Information Systems and marketing

Ilias O. Pappas [a,b,*], Arch G. Woodside [c]

[a] University of Agder, Campus Kristiansand, Universitetsveien 25, 4630 Kristiansand, Norway
[b] Norwegian University of Science and Technology (NTNU), Sem Sælands vei 7-9, 7491, Trondheim, Norway
[c] Boston College, 4801 Wigley Road NE, Marietta, GA 30066, USA

ARTICLE INFO

Keywords:
Asymmetric analysis
Best practices
Calibration
Configurations
fsQCA
Fuzzy sets
Qualitative comparative analysis
Set-theoretic methods

ABSTRACT

The increasing interest in fuzzy-set Qualitative Comparative Analysis (fsQCA) in Information Systems and marketing raises the need for a tutorial paper that discusses the basic concepts and principles of the method, provide answers to typical questions that editors, reviewers, and authors would have when dealing with a new tool of analysis, and practically guide researchers on how to employ fsQCA. This article helps the reader to gain richer information from their data and understand the importance of avoiding shallow information-from-data reporting. To this end, it proposes a different research paradigm that includes asymmetric, configurational-focused case-outcome theory construction and somewhat precise outcome testing. This article offers a detailed step-by-step guide on how to employ fsQCA by using as an example an already published study. We analyze the same dataset and present all the details in each step of the analysis to guide the reader onto how to employ fsQCA. The article discusses differences between fsQCA and variance-based approaches and compares fsQCA with those from structured equation modelling. Finally, the article offers a summary of thresholds and guidelines for practice, along with a discussion on how existing papers that employ variance-based methods are extendable and complemented through fsQCA.

1. Introduction

"Scientists' tools are not neutral" (Gigerenzer, 1991, p. 264): both symmetric (e.g., correlation and multiple regression analysis) and asymmetric (i.e., individual case outcome forecasts) data analysis tools express theoretical stances as well as analytical procedures (Woodside, 2019). Qualitative comparative analysis (QCA) is an asymmetric data analysis technique that combines the logic and empirical intensity of qualitative approaches that are rich in contextual information, with quantitative methods that deal with large numbers of cases and are more generalizable (Ragin, 1987) than symmetric theory and tools. This ability for bringing together basic concepts from both qualitative and quantitative techniques of analysis differs substantially from traditional methods of quantitative analysis that are often variance-based and employ null hypothesis significance testing (NHST). QCA can identify logically simplified statements that describe different combinations (or configurations) of conditions indicating a specific outcome (Ragin, 2008b). A configuration is a specific set of causal variables with a synergetic nature, that serves as a screen indicating an observed outcome or an outcome of interest. QCA has three main variations: crisp set QCA (csQCA), multi-value QCA (mvQCA), and fuzzy-set QCA (fsQCA).

In configurational approaches, the conditions that indicate an outcome, such as user behavior or experience in IS and marketing research, are regarded as configurations of interrelated structures, instead of entities that are examined in isolation (Delery & Doty, 1996; Fiss, 2007). Analyzing alternative configurations enables systemic and holistic views of IS and marketing environments. QCA is useful for inductive, deductive, and abductive (Park, Fiss, & El Sawy, 2020; Saridakis, Angelidou, & Woodside, 2020) reasoning and for remarkably useful theory building, theory elaboration, or theory testing (Greckhamer, Misangyi, & Fiss, 2013; Misangyi et al., 2017). The popularity of QCA and its variations is increasing in different fields such as e-business (Pappas, Kourouthanassis, Giannakos, & Chrissikopoulos, 2016), social media (Pappas, Papavlasopoulou, Mikalef, & Giannakos, 2020), information systems (Liu, Mezei, Kostakos, & Li, 2017; Park & Mithas, 2020; Park et al., 2020), education (Nistor, Stanciu, Lerche, & Kiel, 2019; Pappas, Giannakos, Jaccheri, & Sampson, 2017), and learning analytics and multimodal data (Papamitsiou, Economides, Pappas, & Giannakos,

* Corresponding author at: University of Agder, Norway.
E-mail addresses: ilias.pappas@uia.no, ilpappas@ntnu.no (I.O. Pappas), arch.woodside@bc.edu (A.G. Woodside).

https://doi.org/10.1016/j.ijinfomgt.2021.102310
Received 10 September 2020; Received in revised form 3 January 2021; Accepted 3 January 2021
Available online 5 February 2021

2018; Papamitsiou, Pappas, Sharma, & Giannakos, 2020). At the same time, standards of good practice in research have been published (Woodside, 2016a, 2016b) along with more specific ones that offer a comprehensive and easily accessible "code of good practice" for QCA (Schneider & Wagemann, 2010), including textbooks that describe in detail the method in great detail (Rihoux & Ragin, 2009).

QCA studies are designed to combine techniques from qualitative and quantitative approaches, making these studies inherently mixed technique applications (Ordanini, Parasuraman, & Rubera, 2014; Teddlie & Tashakkori, 2009), that take the best attributes from both worlds. Qualitative inductive reasoning with data being analyzed "by case'' and not "by variable'' (Ragin, 2000), is combined with quantitative empirical testing, as sufficient and necessary conditions identify outcomes through statistical methods (Longest & Vaisey, 2008; Ordanini et al., 2014). In most cases, QCA are useful in quantitative studies, as it allows the researcher to get a deep view of their data through a quantitative analysis that has also several characteristics of qualitative analysis. Case studies focus on describing, explaining, and forecasting, single and combinatorial conditional antecedents on outcomes while variable studies focus on the similarities of variances of two or more variables. A "condition" is a point or interval range of antecedent or outcome; a "variable" characteristic varies. Here are few examples of conditions versus variables: "Male" is a condition; "gender" is a variable. "Swedish" is a condition; "nationality" is a variable. "Expert" is a condition; "expertise" is a variable.

The present tutorial contributes by extending current works as it i) *exemplifies* the application of fsQCA, ii) argues for the need to perform *contrarian case analysis*, and iii) describes how to perform *predictive validity* of the findings. To this end, we map the recommended steps for the above three analyses in two flowcharts that can be found as Appendices in this article. Further, we compare fsQCA with PLS-SE and discuss conceptual differences among the two methods, and finally we summarize some of the frequently used thresholds in fsQCA. This tutorial focuses on quantitative research and offers a step-by-step guide, in an article format on how to employ fsQCA. We focus on fsQCA as this tool can address several shortcomings of symmetric-based analysis. Additional details are provided in the next section, while we offer suggestions on how existing works can be extended by employing fsQCA based on their findings (Appendix D). Our goal is to make fsQCA easy to apply by the scholarly community. Thus, we use as an example a study available in the literature (i.e., (Pappas et al., 2016)) and we offer all the details on how to perform the analysis, that are not included in a typical research article.

This tutorial is structured as follows. First, we present an introduction into the different types of QCA and how they differ. Next, Section 3 presents related studies that have performed fsQCA in IS and Marketing areas. Section 4 introduces complexity theory and how embracing it with fsQCA can take us beyond symmetric tests. Section 5 presents the main benefits and limitations of fsQCA. Next, Section 6 presents a detailed step-by-step guide on how to perform fsQCA while giving the basic steps that should be followed in the analysis, including screenshots from the software. For this a dataset from an already published paper is used. Section 7 presents a comparison between fsQCA and PLS-SEM findings. Section 8 concludes this tutorial.

2. Types of qualitative comparative analysis (QCA)

2.1. CsQCA and mvQCA

CsQCA is the first variation of QCA. It is a tool created to deal with complex sets of binary data (Ragin, 1987). The goal of QCA is to explain causality in complex real life phenomena through "multiple-conjunctural causation, which refers to "nonlinear, nonadditive, non-probabilistic conception that rejects any form of permanent causality and that stresses equifinality (different paths can lead to the same outcome), complex combinations of conditions and diversity" (Berg-Schlosser & De Meur, 2009). QCA uses Boolean algebra and Boolean minimization algorithms to capture patterns of multiple-conjunctural causation and to simplify complex data structures in a logical and holistic manner Ragin (1987). The use of Boolean algebra means that QCA has as input binary data (0 or 1), and uses logical operations for the procedure, thus it is very important to dichotomize the variables in a useful and meaningful manner.

An extension of QCA is mvQCA, which treats variables as multi-valued instead of dichotomous (Cronqvist, 2004). MvQCA retains the idea of performing a synthesis of the dataset and cases with the same value on the outcome variable are explained by a solution, which contains combinations of variables that explain a number of cases with the outcome (Cronqvist & Berg-Schlosser, 2009). Since the method was introduced, a discussion emerged on the potential of mvQCA and its usefulness along with csQCA and fsQCA (Thiem, 2013; Vink & Van Vliet, 2009; Vink & Vliet, 2013), with its set-theoretic status being unclear. MvQCA has remained underutilized (Thiem & Dusa, 2013), compared to the other two variations of QCA (i.e., csQCA and fsQCA).

2.2. FsQCA

FsQCA addresses an important limitation of csQCA, the fact that variables are binary, thus restricting the analysis as it cannot fully capture the complexity in cases that naturally vary by level or degree (Ragin, 2000; Rihoux & Ragin, 2009). This restriction of csQCA is likely an important reason that QCA has not been widely adopted in multiple contexts, including IS and marketing research. FsQCA extends csQCA by integrating fuzzy-sets and fuzzy-logic principles with QCA principles (Ragin, 2000; Rihoux & Ragin, 2009), which offers for a more realistic approach since variables can get all the values within the range of 0–1. FsQCA is able to overcome several limitations of both csQCA and mvQCA, and has received increased attention recently (Fiss, 2011; Ordanini et al., 2014; Pappas et al., 2016; Woodside, 2014), because, when applied together with complexity theory, it provides the opportunity to gain deeper and richer insight into data.

2.3. FsQCA and cluster analysis

Case-based techniques, such as fsQCA and cluster analysis, have been employed as a way of moving beyond variance-based methods (Cooper & Glaesser, 2011). These two techniques have similarities as they both employ multidimensional spaces and often people ask how fsQCA differs from cluster analysis and why do we need it. A main difference between the two methods is the kind of research questions they are able to address (Greckhamer, Furnari, Fiss, & Aguilera, 2018). Specifically, cluster analysis answers questions such as which cases are more similar to each other, while fsQCA can identify the different configurations that constitute sufficient and/or necessary conditions for the outcome of interest (Greckhamer et al., 2018; Ordanini et al., 2014). Depending on the focus of the study the researcher should choose the most appropriate method. Their differences stem from the fact that *"QCA addresses the positioning of cases in [multidimensional] spaces via set theoretic operations while cluster analysis relies on geometric distance measures and concepts of variance minimization"* (Cooper & Glaesser, 2011). To this end, prior studies compare fsQCA with cluster analysis and show how fsQCA can handle causal complexity with fine-grained level data (Fiss, 2011), or how it can identify more solutions compared to cluster analysis (Ordanini et al., 2014). A discussion exists in the literature regarding QCA and cluster analysis (Greckhamer et al., 2018; Miller, 2018), and both approaches have differences making them suitable for different types of studies.

3. Adoption of fsQCA in relevant studies

Configurational approaches are becoming more popular over the past few years in different areas, with fsQCA playing a large part in this

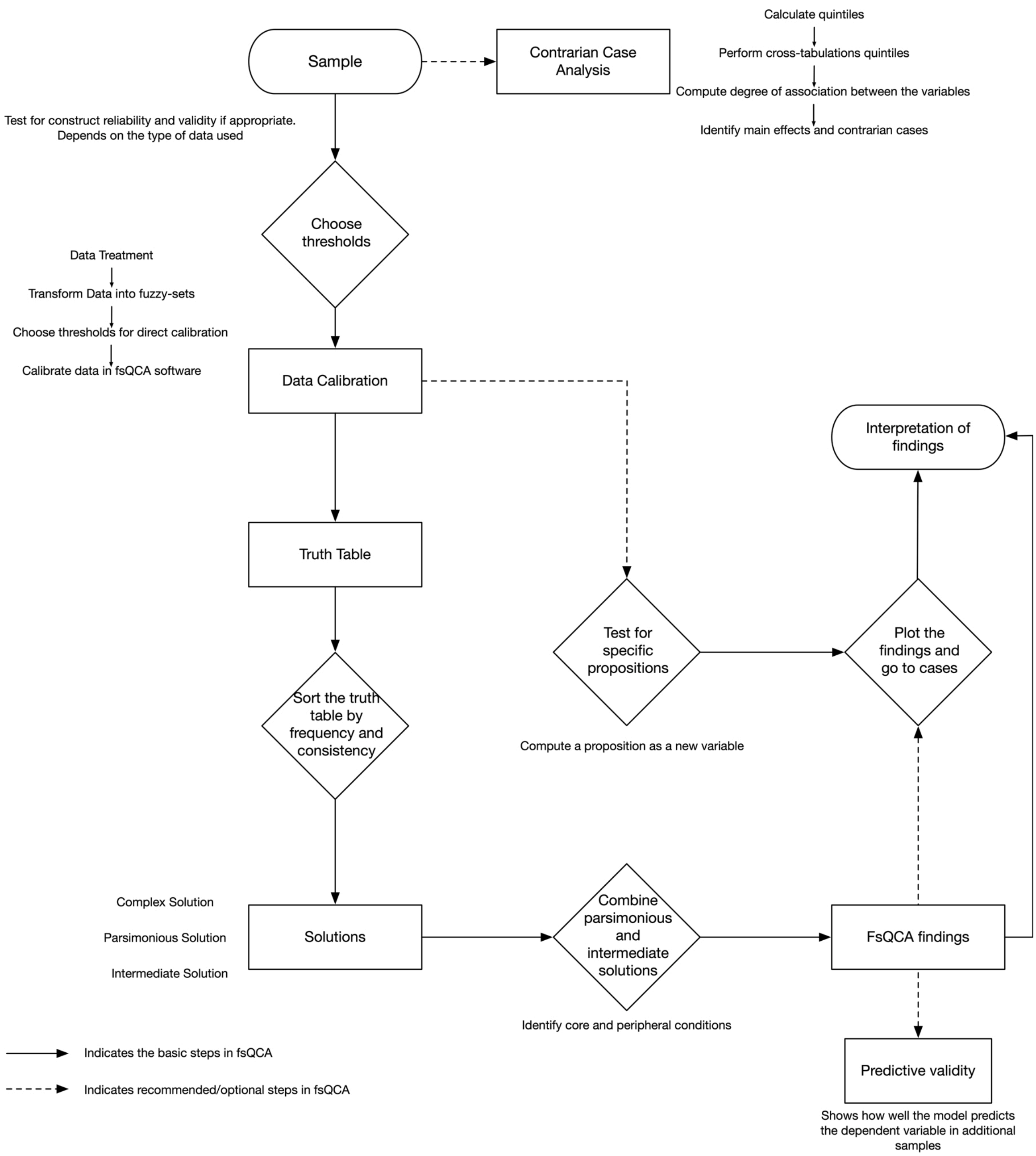
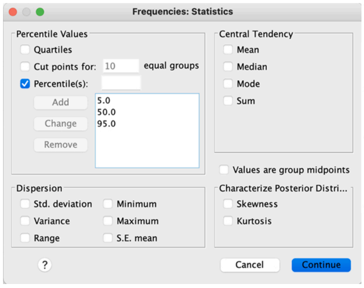


Fig. 1. Basic steps in fsQCA.

as most studies will prefer fuzzy-set over crisp-set and multi-value QCA (Thiem & Dusa, 2013). In detail, fsQCA has been employed in information systems (Fedorowicz, Sawyer, & Tomasino, 2018; Liu et al., 2017), online business and marketing (Pappas et al., 2016; Pappas, 2018; Woodside, 2017), consumer psychology (Schmitt, Grawe, & Woodside, 2017), strategy and organizational research (Fiss, 2011; Greckhamer et al., 2018), education (Pappas, Giannakos et al., 2017; Plewa, Ho, Conduit, & Karpen, 2016), data science (Vatrapu, Mukkamala, Hussain, & Flesch, 2016) and learning analytics (Papamitsiou et al., 2018; Sergis, Sampson, & Giannakos, 2018). This tutorial aims to increase the adoption of fsQCA in IS and marketing studies following the call for more empirical work in the area (El Sawy, Malhotra, Park, & Pavlou, 2010; Fiss, 2011; Woodside, 2014, 2017).

FsQCA is useful for both inductive and deductive reasoning for theory building, elaboration, and testing (Greckhamer et al., 2018; Park et al., 2020). Researchers can either explore all the possible solutions that explain the outcome of interest [e.g., (Pappas et al., 2016)] or test specific models and relations (Pappas, 2018), based on theory or previous findings. Nonetheless, it is possible to perform both, first to identify all possible solutions and then test specific propositions (Pappas, 2018; Pappas et al., 2020). This analysis allows to identify specific cases in the sample, who and how many users, that verify the specific propositions, as well as indicate other alternative models that can explain high score of same outcome. With this knowledge, the researcher can go back to the cases and use contextual information, not including in the analysis, to further explain and discuss the findings. In comparison, a typical variance-based analysis would identify a single best solution, thus limiting the results (Woodside, 2013, 2016b). Nonetheless, studies can compare the findings between different data analysis techniques to describe how different stories are hidden in the same dataset, while it is recommended to combine fsQCA with other data analysis techniques if possible (Schneider & Wagemann, 2010).



Statistics

		Msg_Qual	Ben_Prsnl	Prsnl_Qual	Int_Purchase	Strong_Neg	Weak_Neg	Strong_Pos	Weak_Pos
N	Valid	582	582	582	582	582	582	582	582
	Missing	0	0	0	0	0	0	0	0
Percentiles	5	2.4286	2.6000	2.3333	2.0000	1.0000	1.0000	1.2000	1.0000
	50	4.4286	5.2000	4.6667	4.3333	2.4000	2.0000	3.8000	3.0000
	95	6.1429	6.6000	6.6667	6.6667	4.8000	4.3333	5.8000	5.5000

Fig. 2. Compute thresholds using percentiles.

4. Complexity and configuration theories

Relations among variables are naturally complex, sometimes non-linear, and sudden changes can cause different results and outcomes (Urry, 2005). Variance-based approaches assume that relations among variables are linear, and one way to overcome this is to examine complex phenomena as clusters of interrelated conditions (Woodside, 2017). This offers a step towards a holistic and simultaneous understanding of the patterns these conditions create, by employing a configuration theory approach (El Sawy et al., 2010).

As a destination may usually be reached through different routes, an outcome may occur through different ways, thus explained by different combinations of antecedent conditions. Complexity theory and configuration theories have inherent the principle of *equifinality*, which is the premise that multiple combinations of antecedent conditions are equally effective (Fiss, 2007; Von Bertalanffy, 1968; Woodside, 2014). Numerous factors can influence user experience with any information system and, in general, different combinations of these factors can explain their adoption or use, as well as different levels of the same factors. This means that not all factors (or antecedents) are needed to explain adoption or usage, and it is likely that some of them when combined, can be sufficient to explain high adoption or usage. Nonetheless, in some cases, a factor can be indispensable for high adoption or usage to occur.

Configuration theories are based on the principle of *causal asymmetry*, based on which a condition (or a combination of conditions) that explains the presence of an outcome can be different from the conditions that lead to the absence of the same outcome (Fiss, 2011; Ragin, 2008b). For instance, high perceived usefulness can lead to high intention to use a system, while the low perceived usefulness may not lead to low intention to use a system, typically due to the existence of other conditions. Although such an assumption seems common, when we use variance-based approaches (e.g., correlation, regression) the findings imply that the relation between two variables is symmetrical (high perceived usefulness–high intention to use; low perceived usefulness–low intention to use). In set theory terms, the presence of perceived usefulness (i.e., a condition) may lead to high intention to use, suggesting sufficiency between them. However, high intention to use is very much likely to exist even when perceived usefulness is absent, suggesting that the presence of perceived usefulness is a sufficient but unnecessary condition for intention to use a system. Further, in a different context and when other conditions exist (e.g., high perceived benefits) perceived usefulness may be necessary but insufficient condition for intention to use to occur. Also, sometimes high perceived usefulness may lead to high intention to use only when a third condition is present or absent (e.g., high or low/medium perceived ease of use).

As fsQCA is based on fuzzy-sets, the tool enables capturing conditions that are (1) sufficient or necessary to explain the outcome and (2) insufficient on their own but are necessary parts of solutions that can explain the result. These are called INUS conditions; insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result (Mackie, 1965). Such conditions may be present or absent in a solution, or they may be conditions for which we "do not care". The "do not care" situation indicates that the outcome may either be present

fsqca

msg_qual	ben_prsnl	prsnl_qual	int_purchase	strong_neg	weak_neg	strong_pos	weak_pos	msg_qual_c	ben_prsnl_c	prsnl_qual_c	int_purchase_c	weak_neg_c	weak_pos_c	strong_neg_c	strong_pos_c
6.42857	6.6	6	4.33333	0.5	3.16667	0.5	2	0.97	0.98	0.95	0.62	0.22	0.05	0.5	0.5
4.85714	5.8	4.66667	4.66667	0.05	2.16667	0.71	4.75	0.78	0.94	0.73	0.73	0.06	0.75	0.05	0.71
5	4.6	4	5	0.01	1	0.04	1.5	0.82	0.71	0.5	0.82	0.01	0.02	0.01	0.04
3.42857	2.6	7	4	0.5	5	0.5	4	0.3	0.11	0.99	0.5	0.82	0.5	0.5	0.5
6.14286	6.4	6	5.66667	0.02	1.16667	0.94	6	0.96	0.97	0.95	0.92	0.01	0.95	0.02	0.94
3.42857	4.8	2.33333	3	0.11	1.16667	0.03	1.5	0.3	0.77	0.08	0.18	0.01	0.02	0.11	0.03
5.42857	5.4	5.66667	5.33333	0.08	1.83333	0.89	5	0.89	0.89	0.92	0.88	0.04	0.82	0.08	0.89
4.57143	5.4	5.33333	4.66667	0.14	2.66667	0.71	3.75	0.7	0.89	0.88	0.73	0.12	0.41	0.14	0.71
4.57143	6.6	5.33333	5.66667	0.14	2.16667	0.95	6.25	0.7	0.98	0.88	0.92	0.06	0.97	0.14	0.95
5.57143	5.2	4	2.33333	0.02	1.5	0.08	1.25	0.91	0.86	0.5	0.08	0.02	0.02	0.02	0.08
3.71429	6	4	4.33333	0.43	1.66667	0.04	1	0.39	0.95	0.5	0.62	0.03	0.01	0.43	0.04
4.57143	3.6	4.66667	3.66667	0.02	1	0.02	1	0.7	0.35	0.73	0.38	0.01	0.01	0.02	0.02
4.57143	5.6	3.33333	5	0.43	3	0.5	3.5	0.7	0.92	0.27	0.82	0.18	0.32	0.43	0.5
5	5	5	5	0.02	1.33333	0.02	1	0.82	0.82	0.82	0.82	0.02	0.01	0.02	0.02
6	7	7	7	0.11	2.33333	0.18	4	0.95	0.99	0.99	0.99	0.08	0.5	0.11	0.18
4.14286	5.8	4	4	0.03	1	0.57	3	0.55	0.94	0.5	0.5	0.01	0.18	0.03	0.57
3.85714	5.4	4.33333	4.66667	0.08	1.83333	0.5	2.5	0.45	0.89	0.62	0.73	0.04	0.1	0.08	0.5
5.42857	5.2	5	5.66667	0.18	2.83333	0.43	3	0.89	0.86	0.82	0.92	0.15	0.18	0.18	0.43
4.85714	6.4	5.66667	7	0.02	1	0.11	1.5	0.78	0.97	0.92	0.99	0.01	0.02	0.02	0.11
3.57143	3.8	5	3.66667	0.01	1.16667	0.08	1.75	0.34	0.43	0.82	0.38	0.01	0.03	0.01	0.08
3.57143	3.8	5	3.66667	0.01	1.16667	0.08	1.75	0.34	0.43	0.82	0.38	0.01	0.03	0.01	0.08
3.57143	3.2	3.33333	4	0.18	3	0.35	4.25	0.34	0.23	0.27	0.5	0.18	0.59	0.18	0.35
4.71429	4.4	4.66667	5	0.03	1.16667	0.02	1	0.74	0.65	0.73	0.82	0.01	0.01	0.03	0.02
4	3.4	5	3.33333	0.01	1	0.05	1.25	0.5	0.29	0.82	0.27	0.01	0.02	0.01	0.05
4.85714	6.2	4	5.66667	0.01	2	0.77	4.75	0.78	0.96	0.5	0.92	0.05	0.75	0.01	0.77

Fig. 3. The dataset after calibration.

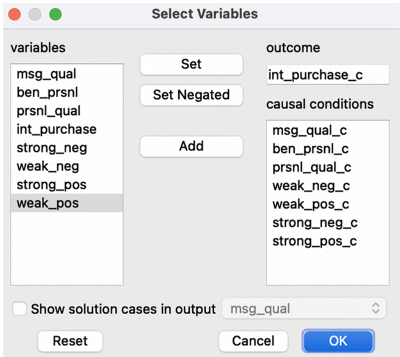


Fig. 4. Selecting variables and outcome for fsQCA.

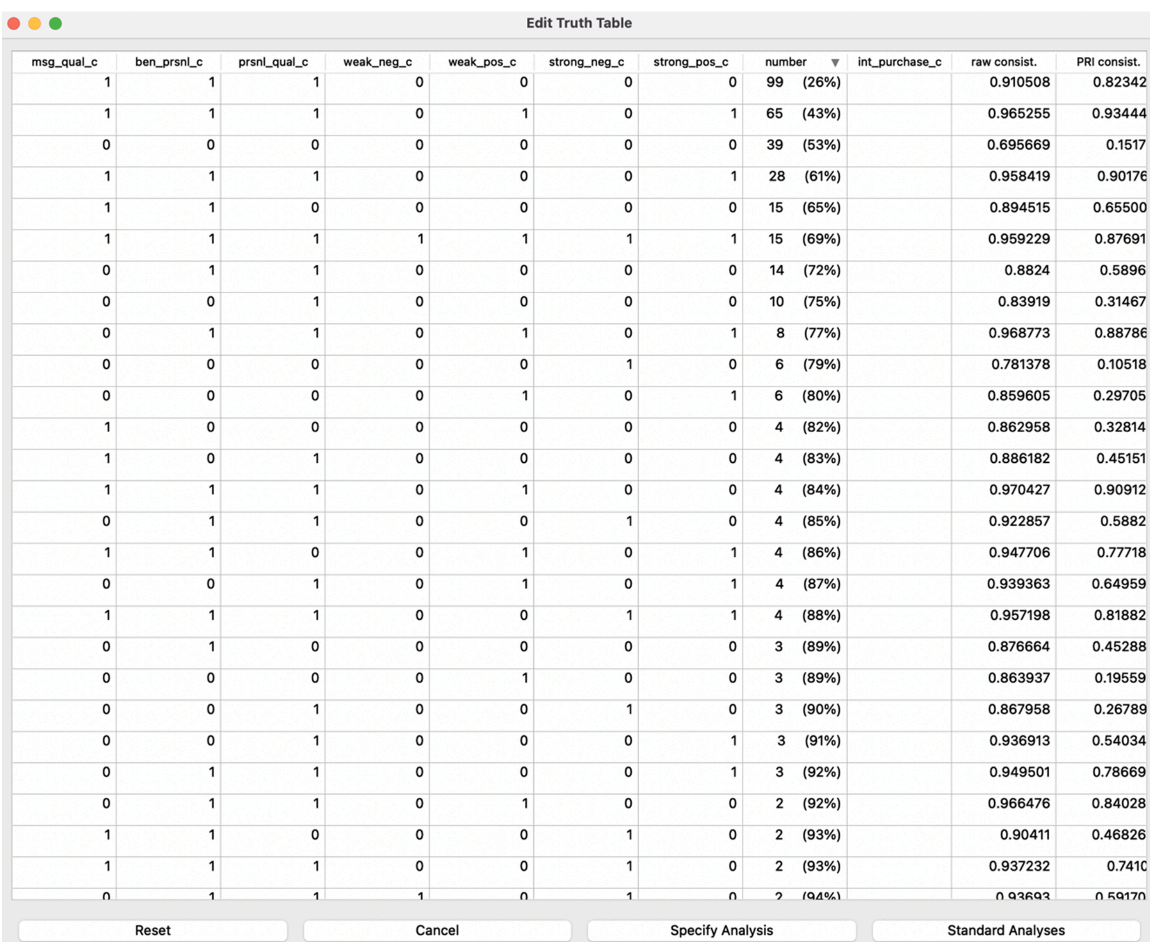
or absent and it does not play a role in a specific configuration. Necessary and sufficient conditions may be present (or absent) as core and peripheral elements. Core elements indicate a strong causal relationship with the outcome, and peripheral elements indicate a weaker relationship (Fiss, 2011). Thus, using fsQCA, researchers can identify which conditions are indispensable (or not needed) for an outcome to occur, and which combinations of conditions are more (or less) important than others

5. The benefits of fsQCA – why to use it?

The use of fsQCA can offer several benefits, compared to traditional methods of analysis. To capture combinations of conditions that are sufficient for an outcome to occur, fsQCA uses both qualitative and quantitative assessments and computes the degree in which a case belongs to a set (Ragin, 2000; Rihoux & Ragin, 2009), thus creating a bridge between qualitative and quantitative methods. FsQCA uses calibrated measures, as data are transformed into the [0, 1] range. Calibration is common in natural sciences but not so much in social sciences and can be used to satisfy qualitative researchers' in interpreting relevant and irrelevant variation as well as quantitative researchers' in precisely placing cases relative to one another (Ragin, 2008b; Vis, 2012).

In line with the focus of this tutorial on quantitative methods, the main benefits of fsQCA occur when compared to typical variance-based methods and the limitations that the latter have (El Sawy et al., 2010; Liu et al., 2017; Woodside, 2013, 2014). In general, variance-based methods examine variables in a competing environment as they compute the net effect between variables in a model, while fsQCA focuses on the complex and asymmetric relations between the outcome of interest and its antecedents. For example, in typical IS/IT adoption studies, following the behavioral science paradigm, variables that are considered as control variables (e.g., gender, experience) can be part of the solutions and combine in ways that explain different parts of the sample. An outcome may be the result of a variety of combinations, and each combination contributes independently to it. In addition, as we seek to design systems that take into account "all users" (i.e., users with different requirements), we need methods that enable researchers to compute multiple solutions for multiple type of users and not the vast majority explained by the best solution of a regression analysis.

FsQCA is employable on different sample sizes ranging from very small (<50 cases) to very large (thousands of cases). The sample size



msg_qual_c	ben_prsnl_c	prsnl_qual_c	weak_neg_c	weak_pos_c	strong_neg_c	strong_pos_c	number	int_purchase_c	raw consist.	PRI consist.
1	1	1	0	0	0	0	99 (26%)		0.910508	0.82342
1	1	1	0	1	0	1	65 (43%)		0.965255	0.93444
0	0	0	0	0	0	0	39 (53%)		0.695669	0.1517
1	1	1	0	0	0	1	28 (61%)		0.958419	0.90176
1	1	0	0	0	0	0	15 (65%)		0.894515	0.65500
1	1	1	1	1	1	1	15 (69%)		0.959229	0.87691
0	1	1	0	0	0	0	14 (72%)		0.8824	0.5896
0	0	1	0	0	0	0	10 (75%)		0.83919	0.31467
0	1	1	0	1	0	1	8 (77%)		0.968773	0.88786
0	0	0	0	0	1	0	6 (79%)		0.781378	0.10518
0	0	0	0	1	0	1	6 (80%)		0.859605	0.29705
1	0	0	0	0	0	0	4 (82%)		0.862958	0.32814
1	0	1	0	0	0	0	4 (83%)		0.886182	0.45151
1	1	1	0	1	0	0	4 (84%)		0.970427	0.90912
0	1	1	0	0	1	0	4 (85%)		0.922857	0.5882
1	1	0	0	1	0	1	4 (86%)		0.947706	0.77718
0	0	1	0	1	0	1	4 (87%)		0.939363	0.64959
1	1	1	0	0	1	1	4 (88%)		0.957198	0.81882
0	1	0	0	0	0	0	3 (89%)		0.876664	0.45288
0	0	0	0	1	0	0	3 (89%)		0.863937	0.19559
0	0	1	0	0	1	0	3 (90%)		0.867958	0.26789
0	0	1	0	0	0	1	3 (91%)		0.936913	0.54034
0	1	1	0	0	0	1	3 (92%)		0.949501	0.78669
0	1	1	0	1	0	0	2 (92%)		0.966476	0.84028
1	1	0	0	0	1	0	2 (93%)		0.90411	0.46826
1	1	1	0	0	1	0	2 (93%)		0.937232	0.7410
0	1	1	1	0	1	0	2 (94%)		0.93693	0.59170

Fig. 5. Truth table in fsQCA.

Edit Truth Table

msg_qual_c	ben_prsnl_c	prsnl_qual_c	weak_neg_c	weak_pos_c	strong_neg_c	strong_pos_c	number	int_purchase_c	raw consist. ▼	PRI consist.
1	1	1	0	1	0	0	4	1	0.970427	0.909128
0	1	1	0	1	0	1	8	1	0.968773	0.887861
1	1	1	0	1	0	1	65	1	0.965255	0.934447
1	1	1	1	1	1	1	15	1	0.959229	0.876914
1	1	1	0	0	0	1	28	1	0.958419	0.901761
1	1	1	0	0	1	1	4	1	0.957198	0.818826
0	1	1	0	0	0	1	3	1	0.949501	0.786694
1	1	0	0	1	0	1	4	1	0.947706	0.777185
0	0	1	0	1	0	1	4	1	0.939363	0.649593
0	0	1	0	0	0	1	3	1	0.936913	0.540342
0	1	1	0	0	1	0	4	1	0.922857	0.58829
1	1	1	0	0	0	0	99	1	0.910508	0.823427
1	1	0	0	0	0	0	15	1	0.894515	0.655007
1	0	1	0	0	0	0	4	1	0.886182	0.451513
0	1	1	0	0	0	0	14	1	0.8824	0.58964
0	1	0	0	0	0	0	3	1	0.876664	0.452882
0	0	1	0	0	1	0	3	0	0.867958	0.267897
0	0	0	0	1	0	0	3	0	0.863937	0.195598
1	0	0	0	0	0	0	4	0	0.862958	0.328144
0	0	0	0	1	0	1	6	0	0.859605	0.297059
0	0	1	0	0	0	0	10	0	0.83919	0.314675
0	0	0	0	0	1	0	6	0	0.781378	0.105182
0	0	0	0	0	0	0	39	0	0.695669	0.15178

Reset | Cancel | Specify Analysis | Standard Analyses

Fig. 6. A sorted truth table in fsQCA based on raw consistency after removing combinations with low frequency.

offers different options to the researcher, either to go back to the cases and interpret them separately, or identify patterns across many cases without returning to the cases (Greckhamer et al., 2013). Further, fsQCA is useful for different types of data (e.g., Likert-scale, clickstreams and multimodal data), as long as the researcher is able to transform them into fuzzy sets. Also, fsQCA is combinable with categorical variables (e.g., gender) which do not have to be transformed into fuzzy sets. In such cases, some variables would be binary (0/1) while some others would cover all values in the [0,1] range.

Data analysis with fsQCA leads to combinations of independent variables that also include variables that are not identified by typical variance based approaches since the capture main affects only (Woodside, 2014). Such variables influence the outcome only for a small number of cases. FsQCA splits the sample into multiple subsets, thus examining multiple combinations of conditions. Each configuration represents only a subset of the sample and an outlier will be present in only some of the possible solutions. In other words, some solutions are likely to explain large parts of the sample, in accordance with a variance-based analysis, while some solutions would explain smaller parts of the sample as they would include cases that would be typically seen as outliers. Thus, in fsQCA thus the representativeness of the sample does not affect all solutions (Fiss, 2011; Liu et al., 2017), making it in a way more robust than variance-based methods as it is not sensitive to outliers. Testing for contrarian cases and examining the distribution of the sample prior to employing fsQCA helps to identify outliers as well as to get an indication there are many cases in the sample that are not explained by the main effects (see details on contrarian case analysis in Section 6.2). Contrarian case analysis is absent typically in research articles.

FsQCA requires that the researcher has accurate workbench knowledge both of the examined variables (conditions and outcome) as well as of the underlying theory and context. Such knowledge is used throughout the analysis; (1) data calibration (i.e., transforming variables into fuzzy sets), (2) simplifying the multiple solutions, (3) interpreting the results. Researchers should make decisions at the different stages based on their knowledge that is typical in qualitative analysis. This action is both a limitation and strength of fsQCA. While it can introduce subjective bias into the study, researchers' own knowledge and understanding of the field and research problem can lead to a richer analysis and understanding of the data. In a traditional mixed-methods approach, a researcher would employ a quantitative study (e.g., analyzing questionnaires, clickstreams, log files) and then employ a qualitative analysis (e.g., interviews with key participants) to gain a richer understanding of their responses and see connections and patterns that are by nature hard to capture. FsQCA is not intended to measure the unique contribution of each variable to the overall observed data; instead, its objective is to identify complex solutions and combinations of independent variables.

Recently, fsQCA has been used to analyze quantitative data (e.g., (Pappas et al., 2016; Vatrapu et al., 2016; Woodside, 2017)), while calibrating qualitative data into fuzzy-sets is also possible (Basurto & Speer, 2012; Henik, 2015). Also, QCA has been employed in mixed-method studies (Cairns, Wistow, & Bambra, 2017) or to bring together quantitative and qualitative data (Kane, Lewis, Williams, & Kahwati, 2014) in different fields. FsQCA provides a novel way of analyzing current datasets, can bring together quantitative or qualitative data, and broaden our methodological approaches and data analyses. Despite its benefits, configurational analysis has limitations that need to be considered when employing fsQCA (Liu et al., 2017; Mendel & Korjani, 2012; Woodside, 2014). Furthermore, best practices have been proposed and should be taken into account when employing QCA (Greckhamer et al., 2018; Schneider & Wagemann, 2010; Woodside,

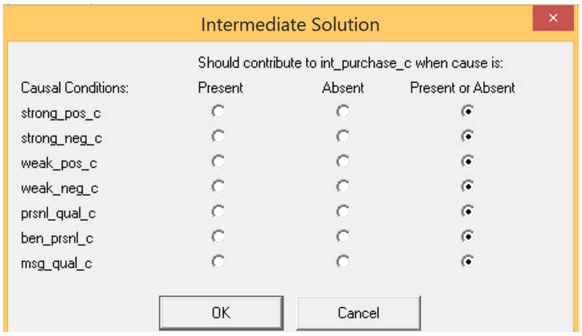


Fig. 7. Setting specific causal conditions as present of absent for the intermediate solution.

2016b). It should be noted that while fsQCA offers increased flexibility to the researcher when it comes to data analysis, and it can be used either for exploratory or confirmatory purposes, researchers should not employ it in a mechanistic way simply following existing guidelines and best practices. Conscious decisions should be made during the analysis, which should be reported to show the validity of the study and enable replicability.

6. How to use fsQCA – Step by step example of employing fsQCA in a typical e-commerce study

For those IS and marketing researchers who are not familiar with fsQCA, this section presents a sample step-by-step analysis of a typical dataset that uses the technique discussed in this article. For this tutorial we use the data and findings from a recently published study (Pappas et al., 2016). We present more details in every step of the process, as well as the guidelines on how to perform fsQCA. Finally, we have visualized the recommended steps for employing fsQCA using fsQCA software in Fig. 1.

6.1. Basic information about the study from the prior published paper

To offer the necessary context for the study, in this section, we offer the basic information regarding the goal and research design. The study examined cognitive and affective perceptions as antecedents of online shopping behavior in personalized e-commerce environments. We used a typical a snowball sampling methodology to recruit participants and controlled for respondents' previous experience with both online shopping and personalized services. Eventually, the sample comprises 582 individuals with experience in online shopping and personalized services. We collected data through a questionnaire built with measures that have been adopted from the literature. Appendix A (as presented in the original study) lists construct definitions, the questionnaire items used to measure each construct, along with descriptive statistics and loadings.

Typical with similar quantitative studies, first we evaluate constructs for reliability and validity. This is a step that should be always performed when it is appropriate, and it is not directly related with the fsQCA analysis as it depends on the type of variables that are used in the study. Construct reliability and validify, as the name implies, refer to the construct itself and not to the method of analysis used to examine relations between constructs. Reliability testing, based on the Cronbach alpha indicator, showed acceptable indices of internal consistency since all constructs exceed the cut-off threshold of 0.70. The AVE for all constructs ranged between 0.55 and 0.84, all correlations were lower than 0.80, and square root AVEs for all constructs were larger than their correlations. The findings in detail for the confirmatory analysis may be found in the original paper (Pappas et al., 2016).

6.2. Contrarian case analysis

Contrarian case analysis is performed outside fsQCA, but we present it here because it can serve as an easy and quick way to examine how many cases in our sample are not explained by main effects, and thus they would not be included in the outcome of a typical variance-based approach (e.g., correlation or regression analysis) (Woodside, 2014, 2016a). Existing studies perform a contrarian case analysis (Pappas

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:
Model: int_purchase_c = f(msg_qual_c, ben_prsnl_c, prsnl_qual_c, weak_neg_c, weak_pos_c, strong_neg_c, strong_pos_c)

 Rows:      23



 Algorithm: Quine-McCluskey
      True: 1

--- COMPLEX SOLUTION ---
frequency cutoff: 3.000000
consistency cutoff: 0.863937

                                                                                      raw       unique
                                                                                    coverage   coverage   consistency
                                                                                   ---------- ---------- ----------
ben_prsnl_c*~weak_neg_c*~weak_pos_c*~strong_neg_c*~strong_pos_c                    0.535961   0.051378   0.837382
~msg_qual_c*prsnl_qual_c*~weak_neg_c*~strong_neg_c*strong_pos_c                    0.261382   0.018454   0.932077
msg_qual_c*ben_prsnl_c*prsnl_qual_c*~weak_neg_c*~strong_neg_c                      0.690433   0.052226   0.917665
msg_qual_c*prsnl_qual_c*~weak_neg_c*~weak_pos_c*~strong_neg_c*~strong_pos_c        0.471245   0.002911   0.895735
~msg_qual_c*prsnl_qual_c*~weak_neg_c*~weak_pos_c*strong_neg_c*~strong_pos_c        0.161339   0.007461   0.877363
msg_qual_c*ben_prsnl_c*prsnl_qual_c*~weak_neg_c*~weak_pos_c*strong_pos_c           0.337572   0.005172   0.956365
msg_qual_c*ben_prsnl_c*~weak_neg_c*weak_pos_c*~strong_neg_c*strong_pos_c           0.337968   0.007178   0.950712
~msg_qual_c*~ben_prsnl_c*~prsnl_qual_c*~weak_neg_c*weak_pos_c*~strong_neg_c*~strong_pos_c   0.118073   0.004663   0.863937
msg_qual_c*ben_prsnl_c*prsnl_qual_c*weak_neg_c*weak_pos_c*strong_neg_c*strong_pos_c   0.133644   0.023597   0.959229
solution coverage: 0.840553
solution consistency: 0.840435
```

Fig. 8. The complex solution.

et al., 2016), however, despite its usefulness, many studies employing fsQCA do not report tests for contrarian cases. Indeed, when examining main relations between two variables, we typically test if a variable positively or negatively affects another variable, which means that most cases in a sample verify this relationship. However, it is likely that the opposite relationship exists for some of the cases in the sample. The existence of such cases can be identified through a contrarian case analysis (Woodside, 2014), since contrarian cases occur regardless of the significance of the main effects.

To perform a contrarian case analysis, first, the sample needed to be divided in order to investigate the relationships among the examined variables. To do so we split the same by using quintiles (i.e., dividing the sample into five equal groups). Other splitting methods of continuous variables, such as median split, should be avoided because this may lead to a reduction of statistical power as well as to false results when the variables are correlated (Fitzsimons, 2008). Next, we performed cross-tabulations across the quintiles, which crosstabs allows us to compute the degree of association between the variables, which suggests a dependence between the two variables and describes main effects between them. The result for any two variables is a 5 × 5 table that presents all combinations for all of the cases in the sample between the two variables (Fig. B4) as shown in Appendix B. The top left and bottom right cases represent the main effects (e.g., degree of association), while the bottom left and top right represent cases not explained by the main effects. The latter are contrarian cases existing in our sample.

All details on how to perform a contrarian case analysis are given in Appendix B. Next, the results for the contrarian case analysis for all variables are presented in Appendix C, as it appears on the original study of Pappas et al. (2016). The findings show the existence of various relationships between the variables, separate from the main effect, supporting the need to perform a configurational analysis.

6.3. Calibration

6.3.1. Data treatment

The most important step in fsQCA is data calibration. Most types of data can be used (e.g., survey responses, clickstreams, user performance data, and physiological data). When a variable or construct is measured with multiple items, we need to compute one value per construct that will be used as input in fsQCA. In other words, for each case (row) in our dataset we need one value for every construct (column). The simplest way is to compute the mean of all the items in order to come up with one single value per case (such as when testing correlations test). Nonetheless, there are more ways to do this while taking into account the individual effect of each item on the construct itself (DiStefano, Zhu, & Mindrila, 2009).

Furthermore, fsQCA does not test for construct reliability and validity as these tests refer to the measures and not the method of analysis. If the constructs used in a study need to be tested for their reliability and validity, then this is done prior to fsQCA analysis, following the traditional methods, and must be reported accordingly.

In fsQCA, different from traditional methods, instead of working with probabilities data are transformed from ordinal or interval scales into degrees of membership in the target set, which shows if and how much a case belongs into a specific set. "In essence, a fuzzy membership score attaches a truth value, not a probability, to a statement" (Ragin, 2008a). For example, the variable intention to purchase can be coded as "high intention to purchase", and we will be looking for the presence or absence of the condition high intention to purchase ("intention to purchase" is the variable; "high intention to purchase" is a condition). Similarly, we code the rest of the variables. The method computes the presence of a condition or its opposite (i.e., negation). The negation of a condition is referred in the literature as the absence of a condition, and the two terms have been used interchangeably based on how the absence is computed (Fiss, 2011; Pappas, 2018; Ragin, 2008b). The term absence has been also used to describe when the condition is irrelevant in a configuration (Nagy et al., 2017; Woodside, 2017), similar to the "do not care" term that is also often used in the literature (Fiss, 2011; Pappas et al., 2016). This distinction is not often addressed or clarified (Pappas, 2018), thus we suggest researchers to clearly define these terms in future works to avoid misunderstandings.

6.3.2. Transform data into fuzzy-sets

In fsQCA we need to calibrate our variables to form fuzzy sets with their values ranging from 0 to 1 (Ragin, 2008b). Consider a fuzzy set as a group, then the values from 0 to 1 define if and at what amount a case belongs to this group. The fact that all values range from 0 to 1 means that a case with a fuzzy membership score of 1 is a *full member* of a fuzzy set (fully in the set), and a case with a membership score of 0 is a *full non-member* of the set (fully out of the set). A membership score of 0.5 is exactly in the middle, thus a case would be both a member of the fuzzy set and a non-member, and is therefore a member of what is known as the *intermediate* set. The intermediate-set point is the value where there is maximum ambiguity as to whether a case is more in or more out of the target set.

Data calibration may be either direct or indirect. In the direct calibration the researcher needs to choose exactly three qualitative breakpoints, which define the level of membership in the fuzzy set for each case (fully in, intermediate, fully out). In the indirect method, the measurements need to be rescaled based on qualitative assessments. The researcher may choose to calibrate a measure differently, depending on what one is investigating. Either method may be chosen, depending on researcher's substantive knowledge of both data and underlying theory (Rihoux & Ragin, 2009). The direct method is recommended and is more common, in which the researcher sets three values corresponding to full-set membership, full-set non-membership, and intermediate-set membership. This can lead to more rigorous studies which are easier to be replicated and validated, since it is clearer on how the thresholds have been chosen.

6.3.3. Choosing thresholds for direct calibration

To calibrate the data, we can choose the values 0.95, 0.50, 0.05 as the three thresholds (or breakpoints), which will transform the data into the log-odds metric with all values being between 0 and 1. We do not use exactly 1 and 0 as breakpoints because the two membership scores would correspond to positive and negative infinity, respectively, for the log of the odds (Ragin, 2008a). To find which values in our dataset correspond to the 0.95, 0.50, and 0.05 we use percentiles. The percentiles allow the calibration of any measure regardless of its original values. In detail, we can compute the 95 %, 50 %, and 5 % of our measures and use these values as the three thresholds in fsQCA software. This can be done easily, for example in SPSS using the *"Percentiles"* function (Frequencies > Statistics > Percentiles) and choosing the desired percentages (Fig. 2). Nonetheless, the thresholds should not be chosen mechanically but the researcher should understand what it means for example that 5 % defines being fully out of the set. This may mean that the threshold should be changed or adjusted, since it is up to the researcher to select the three thresholds. For example, if the data do not have a normal distribution but instead are skewed then the 80 %, 50 %, and 20 % can be set as the thresholds for full-set membership, intermediate-set membership, and full-set non-membership, respectively (Pappas, Mikalef, Giannakos, & Pavlou, 2017). In any case, the choice of thresholds should be justified and reported accordingly, along with a table that presents the original values that correspond to each threshold (Fig. 2).

Especially in the case of the widely used seven-point Likert scales (1=Not at all, 7=Very much), previous studies suggest that the values of 6, 4, and 2 can be used as thresholds (Ordanini et al., 2014; Pappas et al., 2016). Similarly, for a five-point Likert scale the thresholds could be 4,3, and 2. For the example presented in Fig. 2, a seven-point Likert scale was used. We note that, for most variables, the percentiles give the same values as if we had chosen directly the values of 6,4, and 2. However,

```
*********************
*TRUTH TABLE ANALYSIS*
*********************

File:
Model: int_purchase_c = f(msg_qual_c, ben_prsnl_c, prsnl_qual_c, weak_neg_c, weak_pos_c, strong_neg_c, strong_pos_c)

 Rows:      23



 Algorithm: Quine-McCluskey
      True: 1-L

--- PARSIMONIOUS SOLUTION ---
frequency cutoff: 3.000000
consistency cutoff: 0.863937

                                raw       unique
                              coverage   coverage   consistency
                              ---------- ---------- ----------
ben_prsnl_c                   0.924911   0.117819   0.794118
weak_pos_c*~strong_pos_c      0.234760   0.004494   0.887975
msg_qual_c*prsnl_qual_c       0.765295   0.004833   0.896600
prsnl_qual_c*strong_neg_c     0.258556   0.004154   0.871499
prsnl_qual_c*strong_pos_c     0.550176   0.008337   0.922086
solution coverage: 0.958259
solution consistency: 0.773678
```

Fig. 9. The parsimonious solution.

this is not the case for all variables. In detail, Weak_Negative emotions have the lowest scores overall, and the 95 %, 50 %, and 5 % are the values 4.33, 2.00, and 1.00. Since this is a construct measured with a seven-point Likert scale, if we use the 95 % it means that users with scores 4.33 or higher fully belong to the set, which is high values of weak negative emotions. However, this would be an inaccurate representation of those cases, as users that choose 4 or 5 are nearer to the neutral point rather than to the higher point in the scale. Thus, using the thresholds 6, 4, 2 provides a more accurate representation of our sample.

6.3.4. Calibrating the data in fsQCA software

Once we have decided the thresholds, we proceed to the data calibration in fsQCA software (version 3.0 is used here) (Ragin & Davey, 2016). The dataset file needs to be in "comma-separated values" (.csv) format to be able to open it in fsQCA. Calibration is performed by using the *Calibrate* function of the software, which takes as input the variable that will be calibrated and the three breakpoints (from the highest to the lowest values) (Fig. 3). It should be noted that the researchers may use other software for the calibration part, and that it is not mandatory to calibrate all values following a logistic function; instead, other

```
*********************
*TRUTH TABLE ANALYSIS*
*********************

File:
Model: int_purchase_c = f(strong_pos_c, strong_neg_c, weak_pos_c, weak_neg_c, prsnl_qual_c, ben_prsnl_c, msg_qual_c)

 Rows:      22



 Algorithm: Quine-McCluskey
      True: 1
  0 Matrix: 0L
Don't Care: -

--- INTERMEDIATE SOLUTION ---
frequency cutoff: 3.000000
consistency cutoff: 0.863937
Assumptions:

                                                                              raw       unique
                                                                            coverage   coverage   consistency
                                                                            ---------- ---------- ----------
~strong_pos_c*~strong_neg_c*~weak_pos_c*~weak_neg_c*ben_prsnl_c             0.535961   0.051378   0.837382
strong_pos_c*~strong_neg_c*~weak_neg_c*prsnl_qual_c*~msg_qual_c             0.261382   0.018454   0.932077
~strong_neg_c*~weak_neg_c*prsnl_qual_c*ben_prsnl_c*msg_qual_c               0.690433   0.052226   0.917665
~strong_pos_c*strong_neg_c*~weak_pos_c*~weak_neg_c*prsnl_qual_c*~msg_qual_c 0.161339   0.007461   0.877363
~strong_pos_c*~strong_neg_c*~weak_pos_c*~weak_neg_c*prsnl_qual_c*msg_qual_c 0.471245   0.002911   0.895735
strong_pos_c*~strong_neg_c*weak_pos_c*~weak_neg_c*ben_prsnl_c*msg_qual_c    0.337968   0.007178   0.950712
strong_pos_c*~weak_pos_c*~weak_neg_c*prsnl_qual_c*ben_prsnl_c*msg_qual_c    0.337572   0.005172   0.956365
~strong_pos_c*~strong_neg_c*weak_pos_c*~weak_neg_c*~prsnl_qual_c*~ben_prsnl_c*~msg_qual_c 0.118073 0.004663 0.863937
strong_pos_c*strong_neg_c*weak_pos_c*weak_neg_c*prsnl_qual_c*ben_prsnl_c*msg_qual_c 0.133644 0.023597 0.959229
solution coverage: 0.840553
solution consistency: 0.840435
```

Fig. 10. The intermediate solution.

Table 1
FsQCA findings.

	Solution								
Configuration	1	2	3	4	5	6	7	8	9
Cognitive Perceptions									
Quality of Personalization	●	●	●	●	●	●			⊗
Message Quality	⊗	●	●	●	●	⊗		•	⊗
Benefits of Personalization		●	●	●			●	●	⊗
Affective Perceptions									
Strongly Positive	●	●	●		⊗	⊗	⊗	•	⨂
Weakly Positive		⊗	•		⊗	⊗	⊗	•	●
Strongly Negative	⊗		●	⊗	⊗	●	⊗	⊗	⊗
Weakly Negative	⊗	⊗	•	⊗	⊗	⊗	⊗	⊗	⊗
Consistency	0.932	0.956	0.959	0.918	0.896	0.877	0.837	0.950	0.863
Raw Coverage	0.261	0.337	0.133	0.690	0.471	0.161	0.535	0.337	0.118
Unique Coverage	0.018	0.005	0.023	0.052	0.002	0.007	0.051	0.007	0.004
Overall solution consistency	*0.841*								
Overall solution coverage	*0.840*								

Note: Black circles (●) indicate the presence of a condition, and circles with “x” (⊗) indicate its absence. Large circle; core condition, Small circle; peripheral condition, Blank space; “don’t care” condition.

membership functions (linear or non-linear) may be used (Mendel & Korjani, 2012). Besides, the fsQCA software a package for R exists as well (Thiem & Dusa, 2013).

In fsQCA, the cases that are exactly on 0.5 are dropped from the analysis which makes it difficult to analyze the conditions that are set exactly on 0.5 (i.e., intermediate-set membership) (Ragin, 2008b). To overcome this, Fiss (2011) suggests adding a constant of 0.001 to the causal conditions below full membership scores of 1. This can be done by adding 0.001 in all conditions after the calibration has been performed.

Once all variables have been calibrated the dataset includes both versions of each variable (Fig. 3) and we can proceed to the next step, which is running the fuzzy-set algorithm and the generation of the truth table.

In order to run the fsQCA algorithm choose “Analyze > Truth Table Algorithm”, and at this point the researcher has to select the variables that will be included in the analysis (Figure 4). In detail, the causal conditions are the independent variables, and the outcome is the dependent variable. At this point, the researcher may choose to either compute the presence of outcome or negation of the outcome. By clicking *OK,* fsQCA produces the truth table.

The truth table computes all possible configurations (or combinations) that may occur, providing 2^k rows, with k representing the number of outcome predictors, and each row representing every possible combination (Fig. 5). When computing all possible configurations, the frequency is also presented (i.e., the number of observations for each possible combination), while several lines will have a frequency of zero, which means that none of the cases in the sample are explained by them. As the number of variables in an analysis increases, the number of possible configurations increases exponentially (2^k), thus the more variables the more combinations are likely to have a frequency of zero. Including more variables in the analysis might benefit by having also a larger sample, which is common in typical quantitative analysis (e.g., MRA, SEM).

Next, the truth table needs sorting by frequency and consistency (Ragin, 2008b). This is done using the options in the Sort menu. Since frequency describes how many cases in the sample are explained by a configuration, to ensure that a minimum number of cases is obtained for the assessment of the relationships, a frequency threshold is set (i.e., column number) (Fig. 5). A higher frequency threshold means that each configuration refers to more cases in the sample, but as a result will reduce the percentage (i.e., coverage) of the sample that is explained by the retained configurations. On the other hand, a lower frequency threshold increases the coverage of the sample, although each combination refers to fewer cases in the sample. For samples larger than 150 cases the frequency threshold may be set at 3 (or higher), while for smaller samples the threshold may be set at 2 (Fiss, 2011; Ragin, 2008b). As our sample is 582, the threshold is set at 3, and all combinations with smaller frequency are removed from further analysis.

Once removing configurations with low frequency, the truth table should be sorted by “raw consistency” (Fig. 6) At this point a consistency threshold should be set, with the minimum recommended value being 0.75 (Rihoux & Ragin, 2009). A first indicator for choosing the consistency threshold is to identify natural breaking points in the consistency values that have been obtained, however this is not absolute. In detail, in Fig. 6 we note the lowest consistency values of 0.862958, 0.859605, 0.839190, 0.781378, 0.695669. These values indicate that both 0.781378 and 0.839190 may be breaking points and potential frequency thresholds. Thus, the researcher needs to decide which is the appropriate threshold and justify this choice. To aid the researcher, fsQCA software calculates PRI consistency, which stands for “Proportional Reduction in Inconsistency” and is an alternate measure of the consistency of subset relations in social research, and only relevant to fuzzy sets. PRI consistency is used to avoid simultaneous subset relations of configurations in both the outcome and the absence of the outcome (i.e., negation). PRI consistency scores should be high and close to raw consistency scores (e.g., 0.7), while configurations with PRI scores below 0.5 indicate significant inconsistency (Greckhamer et al., 2018). Thus, a PRI consistency threshold should also be used.

Finally, fsQCA software calculates SYM consistency (i.e., Symmetric Consistency), which was developed for fuzzy-sets and can be used when the researcher examines both the presence and negation of the outcome and wants to use the same consistency standard for both analyses (i.e., presence and its negation). In general, most papers do not present the truth table in their analysis, however presenting this can increase the validity of the findings and strengthen the rigorousness of the process. It should be noted that a low consistency threshold leads to the identification of more necessary conditions, reducing type II errors (i.e., false negatives), but increasing type I errors. (i.e., false positives), and vice versa (Dul, 2016).

The last step while working on the truth table is to decide if each combination (i.e., line of the table) explains the outcome or not. Using the consistency thresholds, the researcher has to insert the value of 1 or 0 in the column with the outcome variable. Choosing 1 or 0 defines if a combination explains the outcome or not. Once this is complete, the researcher may proceed to obtain solution sets (command: *Standard Analyses*).

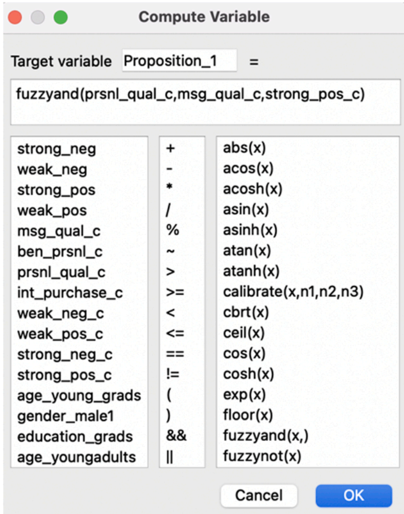


Fig. 11. Computing a specific proposition (model).

Next, the researcher may decide whether a single independent variable should be present or absent at all times in the chosen configurations (Fig. 7), contributing to the intermediate solution (to be explained below). Unless otherwise needed (e.g., based on the theory or literature), it is suggested to choose "Present or Absent" in order to obtain all the possible configurations.

6.4. *Obtaining the configurations/solutions*

FsQCA computes three solution, namely complex solution, parsimonious solution, and intermediate solution. Here, "solution" refers to a combination of configurations that is supported by a high number of cases, where the rule "the combination leads to the outcome" is consistent.

The *complex* solution presents all the possible combinations of conditions when traditional logical operations are applied (Fig. 8). In general, because the number of identified configurations can be very large, the number of complex solutions can be very large and these may include configurations with several terms, making the interpretation of the solutions rather difficult and in most cases impractical. For this reason, they are simplified further into parsimonious and intermediate solution sets.

The *parsimonious* solution set is a simplified version of the complex solution, based on simplifying assumptions, and presents the most important conditions which cannot be left out from any solution (Fig. 9). These are called "core conditions" (Fiss, 2011) and are identified automatically by fsQCA. The major difference between parsimonious and complex solution is that the complex solution excludes counterfactual cases, involving limited simplification, while the parsimonious solution includes any counterfactual combination that can contribute to a logically simpler solution.

Finally, the *intermediate* solution is obtained when performing counterfactual analysis on the complex and parsimonious solutions including only theoretically plausible counterfactuals (Liu et al., 2017; Ragin, 2008b) (Fig. 10). The intermediate solution uses a subset of those simplifying assumptions used to compute the parsimonious solution, which should be consistent with theoretical and empirical knowledge. Based on previous knowledge the researcher may choose if one of the variables should be considered as only present, only absent, or either, in explaining the outcome. By default, either present or absent is computed. Any decisions made regarding the connection between each causal condition and the outcome need to be based on theoretical or substantive knowledge (Fiss, 2011). The intermediate solution is part of the complex solution and includes the parsimonious solution. While core conditions appear in both parsimonious and intermediate solutions, the

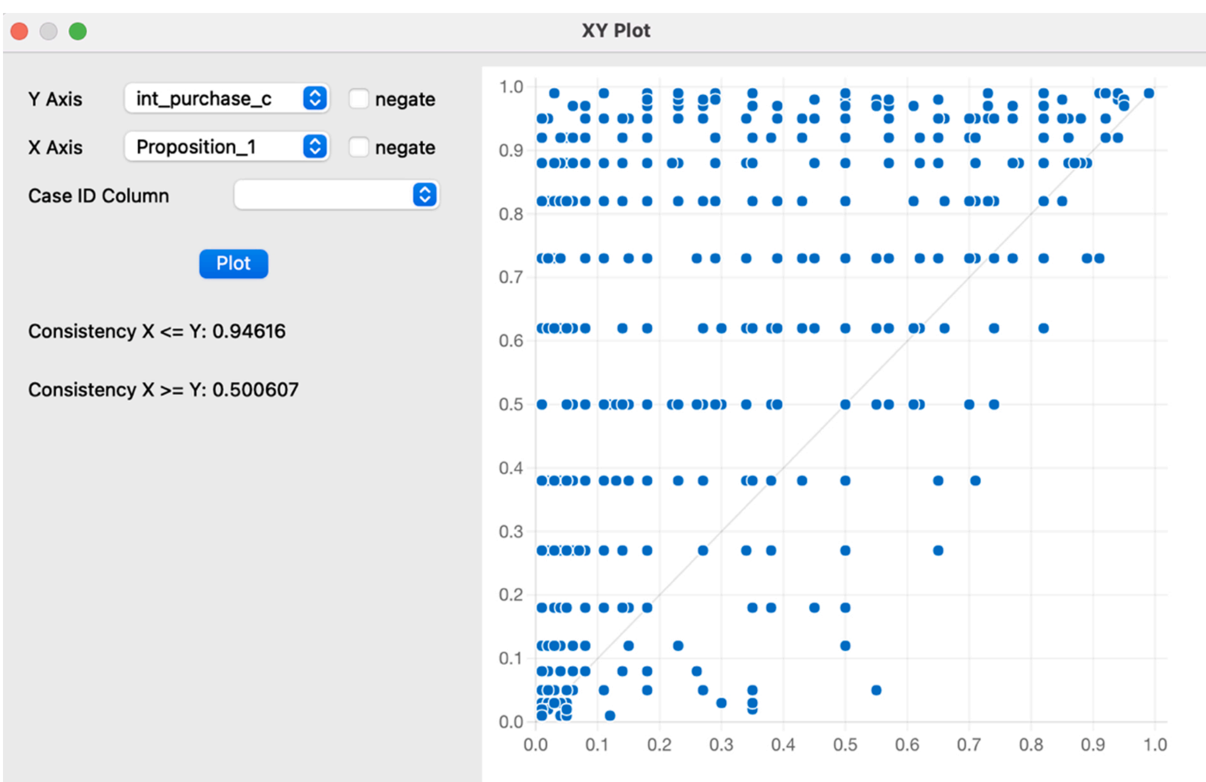


Fig. 12. Plotting a specific proposition (Proposition 1; *high quality of personalization, high message quality, and high strongly positive emotions will have high intention to purchase.*

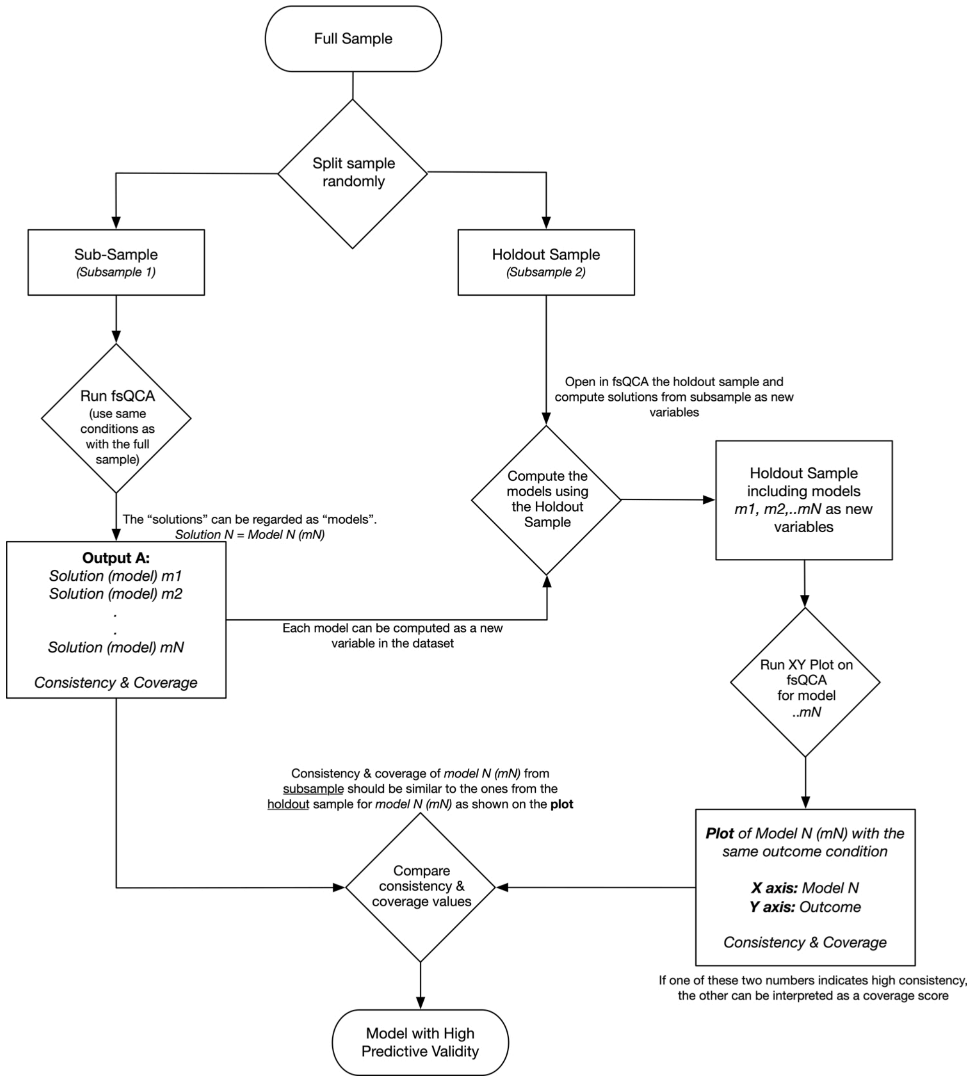


Fig. 13. Basic steps in performing predictive validity.

Table 2
Solutions (models) from the subsample.

Models from subsample	Raw coverage	Unique Coverage	Consistency
BP●MQ●~SP●~SN●~WP●~WN	0.532	0.044	0.877
QP●BP●~MQ●~SP●~WP●~WN	0.309	0.027	0.853
QP●BP●~MQ●~SP●~WP●~WN	0.309	0.027	0.853
QP●BP●SP●~SN●WP●~WN	0.349	0.011	0.957
QP●BP●MQ●SP●SN●WP●WN	0.125	0.018	0.946
QP●BP●MQ●~SN●~WP●~WN	0.559	0.015	0.895
QP●BP●MQ●SP●~SN●~WN	0.459	0.012	0.949
Overall solution consistency	*0.869*		
Overall solution coverage	*0.791*		

QP; Quality of Personalization, BP; Benefits of Personalization, MQ; Message Quality, SP; Strongly Positive Emotions, WP; Weakly Positive Emotions, SN; Strongly Negative Emotions, WN; Weakly Negative. ●; Logical conjunction (AND), ~; Negation (NOT).

conditions that are eliminated in the parsimonious solution and appear only in the intermediate solution are called "*peripheral conditions*" (Fiss, 2011). In other words, since the intermediate solution presents both core and peripheral conditions, and peripheral conditions are removed from the parsimonious solution, an easy way to identify the core conditions is to examine the parsimonious solution since it does not include peripheral conditions. Also, the parsimonious solution is typically smaller than the intermediate. However, it is possible that they could be exactly the same, meaning that no elaboration is useful beyond the parsimonious solution. By including additional conditions in the solution, we increase the complexity in favor of increased consistency. Comparing Figs. 9 and 10, we see that the intermediate solution has a higher consistency than the parsimonious. A more detailed and mathematically justified description of the steps in counterfactual analysis is available in Mendel and Korjani (2012).

6.5. Interpreting and presenting the solutions

FsQCA software provides all three solutions every time. Complex and parsimonious solutions are computed regardless of any simplifying assumptions employed by the researcher (e.g., choosing the presence or absence/negation of a variables) while the intermediate solution depends on these assumptions. While the intermediate solution includes both core and peripheral conditions, we need an easy way to make the distinction that will help us interpret and present the solutions in a better manner.

Combining the parsimonious and intermediate solutions can offer a more detailed and aggregated view of the findings (Fiss, 2011). To do this, the researcher can identify and mark the conditions of the parsimonious solution that also appear in the intermediate solution. This

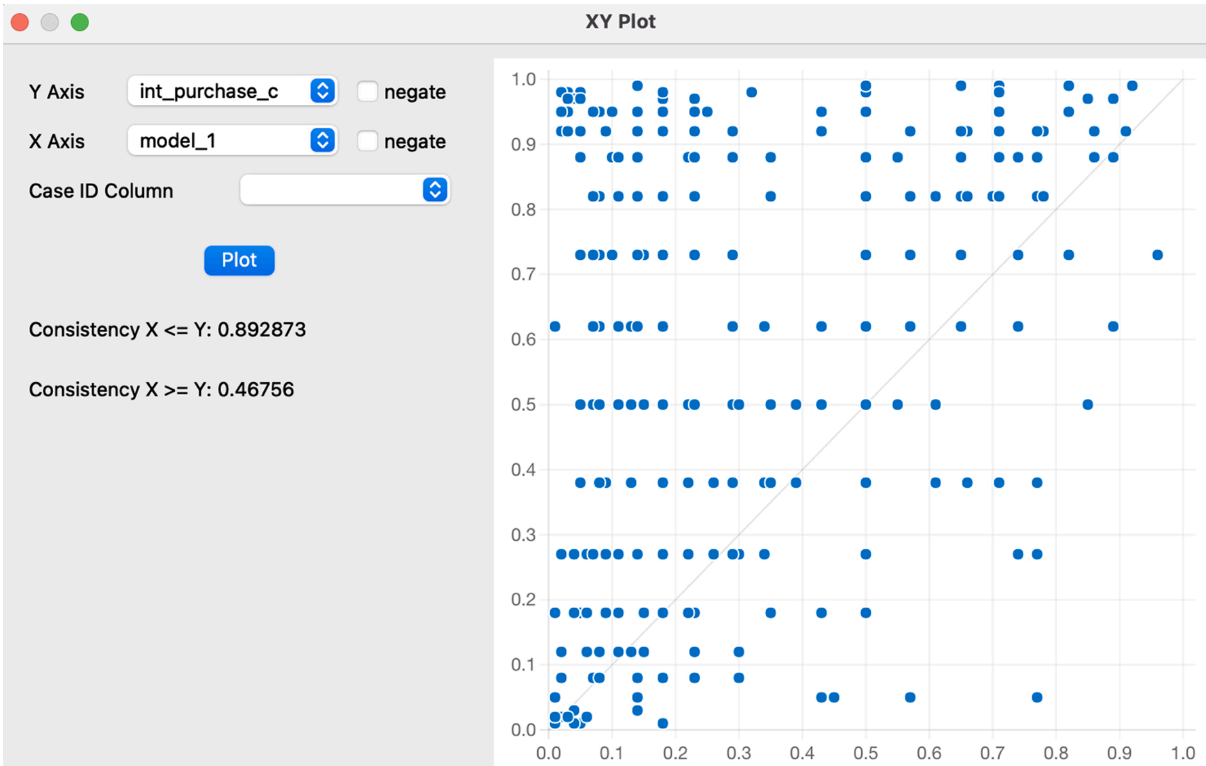


Fig. 14. Fuzzy-Plot of Model 1 (from Table 2) using data from the holdout sample.

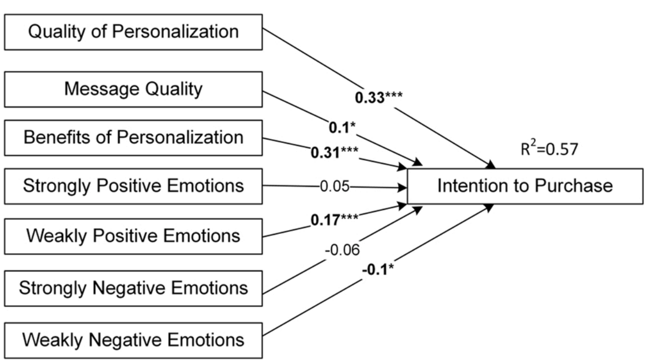


Fig. 15. Findings from PLS-SEM analysis.

practically will lead to an intermediate solution which has highlighted the core conditions, clearly presenting all core and peripheral conditions, allowing for a better interpretation of the findings. Frequently, we may have situations where more than one core condition co-occurs in a given case. Hypothetically, if we have a parsimonious solution of A + BC + BD and an intermediate solution of AcD + BCE + ABF + ABCDf, we report **A**cD + **BCE** + **AB**F + **ABCD**f, with bold characters indicating core conditions. Nonetheless, the researcher may choose to present only the parsimonious solution and focus only on the core conditions that cannot be left out of any solution.

Next, to improve the presentation of the findings we can transform the solutions from fsQCA output (Figs. 9 and 10) into a table that is easier to read (Table 1). Typically, the presence of a condition is indicated with a black circle (●), the absence/negation with a crossed-out circle (⊗), and the "do not care" condition with a blank space (Fiss, 2011). As we mentioned earlier (see Section 6.3.1), the negation of a condition is referred in the literature also as absence, and the two terms have been used interchangeably (Pappas, 2018). The distinction between core and peripheral is made by using large and small circles, respectively. The researcher needs to present the overall solution consistency and the overall solution coverage. The overall coverage describes the extent to which the outcome of interest may be explained by the configurations, and is comparable with the R-square reported on regression-based methods (Woodside, 2013). In our example, the results indicate an overall solution coverage of 0.84, which suggests that a substantial proportion of the outcome is covered by the nine solutions.

The fsQCA findings, as reported in the original article, are readable as follows. For high purchase intentions to occur, solutions 1–3 reflect combinations of the presence and absence of cognitive with affective perceptions. Quality of personalization and strongly positive emotions are core constructs, pointing out the importance of these factors. In detail, the combination of high quality of personalization with strongly positive emotions towards personalized services, with the absence of message quality and both types of negative emotions, lead to high purchase intentions, regardless of the level of benefits of personalization and weakly positive emotions (solution 1). To this end, when all cognitive perceptions are present, in order to achieve high purchase intentions, they may be combined with either (i) strongly positive emotions, with the absence of weakly positive and negative emotions, and regardless of strongly negative emotions (solution 2), or (ii) all types of emotions (solution 3), or (iii) with the absence of negative emotions and regardless of positive ones (solution 4). With the absence of all emotions, high purchase intentions may be achieved with either high personalization and message quality, regardless of their benefits (solution 5), or just by high personalization benefits regardless of its quality (solution 7). Solution 6 combines personalization quality and strongly negative emotions, with the absence of message quality along with the rest of emotions. Personalization benefits play a minor role in this

Table 3
Findings using fsQCA and PLS-SEM on the same sample.

FsQCA findings	PLS-SEM findings
Quality of personalization is present in 6 out of 9 solutions that explain intention to purchase *This can also be read as: High quality of personalization explains intention to purchase in 6 out of 9 solutions.*	Quality of personalization is an important predictor of intention to purchase
Benefits of personalization are present in 5 out of 9 solutions that explain intention to purchase	Benefits of personalization are an important predictor of intention to purchase
Cognitive perceptions are more frequently present than absent when explaining intention to purchase	All cognitive perceptions influence intention to purchase
Affective perceptions are more frequently absent or negated than present when explaining intention to purchase	Weakly positive and negative emotions influence intention to purchase
Strongly positive and negative emotions can be either present or negated in explaining intention to purchase depending on how they combine with the other factors.	Strongly positive and negative emotions have no effect on intention to purchase
Message quality and weakly negative emotions can be either present or negated in explaining intention to purchase depending on how they combine with the other factors.	Message quality and weakly negative emotions have weak effects (0.1) on intention to purchase.
Different levels of weakly negative emotions can contribute to explaining intention to purchase. Weakly negative emotions are present in one solution, and not present in eight solutions.	Weakly negative emotions have a negative effect on intention to purchase
Different levels of weakly positive emotions may contribute to explaining intention to purchase. Weakly positive emotions are present in three solutions, but not present in four solutions.	Weakly positive emotions have a positive effect on intention to purchase
All cognitive perceptions and emotions can be present at the same time and explain intention to purchase	Not all cognitive perceptions and emotions influence intention to purchase.
Positive and negative emotions can coexist in explaining intention to purchase.	Weakly positive and weakly negative emotions may coexist in explaining intention to purchase.
The presence of (weakly positive) emotions can lead to high intention to purchase even when cognitive perceptions are not present.	Weakly negative emotions have a negative effect on intention to purchase
The multiple solutions may refer to different types of users, with different perceptions and needs.	The single solution is the “best” solution that explains the largest percentage of variance.
The results indicate an overall solution coverage of .84, which suggests that a substantial proportion of intention to purchase is covered by the nine solutions	The results indicate an R^2 of 0.57, which means that 57% of the variance of intention to purchase is explained by the model

solution. On the other hand, in solution 8, benefits are an important (core) factor which combined with message quality, and the presence of positive emotions only lead to high purchase intentions. Finally, the same outcome may be achieved by the presence of weakly positive emotions combined with the absence of all other emotions and all cognitive perceptions (solution 9).

6.6. Testing for specific propositions

After obtaining all the possible solutions that can explain the outcome of interest with fsQCA, we can also test for specific propositions and examine for how many cases in the sample these propositions hold true (Pappas, 2018; Pappas et al., 2020). In addition, we can identify specifically which cases are these in the sample, if such information is relevant. This is performed by computing the specific configuration in fsQCA, thus creating a model, and plotting it against the outcome of interest.

Table 4
Conceptual differences between fsQCA and RBMs.

FsQCA	Regression based models (RBMs)
Relationships between variables can be either symmetric or asymmetric, non-linear.	Relations are assumed to be symmetrical and linear.
The conditions that explain high levels of an outcome are not exact opposites of the ones that explain low levels of an outcome (causal asymmetry)	We focus on identifying determinants that explain high levels of an outcome, assuming that the exact opposite will lead to low levels of the same outcome.
Allows for case-based modelling to identify localized effects and also to return to the cases after the analysis for deeper understanding of the results and the sample.	Focuses on the unique contribution of a variable while holding constant the values of all other variables in the equation.
We identify multiple solutions as algorithms that are sufficient or necessary to explain the same outcome (equifinality).	We identify a single best solution (unifinality).
We identify the INUS condition, that is insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result.	We report on the most important conditions that explain an outcome. We can examine direct, indirect, and moderating effects
We explain both main effects and contrarian cases	Findings explain cases that are represented by main effects.
FsQCA transforms variables into scales and gives them a truth value that defines their membership in the set.	RBMs use probabilities to compute estimates that provide the single best solution.
FsQCA computes the coverage, which based on the truth value that they receive, is the actual number of cases explained by the solution.	RBMs, using probabilities, estimate the variance of the model which is presented with the R^2
Solutions are computed based on the coverage value and the existence of cases in the sample that explain a configuration, and consistency is used to define the strength of a relationship, supported by empirical evidence.	Acceptance or rejection is based on the effect size and the p-value is used to determine its significance.
Variables work together and combine with each other to explain the outcome	Variables compete with each other to explain the largest percentage of variance.
We develop propositions that can examine the role of combinations of variables in explaining the outcome and allow case identification in the model.	We develop hypotheses that depict the relation between 2 variables.
Counter-hypothesized relations are expected and can be explained.	Hypothesis of positive or negative effects influences the findings and the acceptance or rejection of hypotheses. It is difficult to explain counter hypothesized relations.

The first step is to compute the configuration as described by the proposition and transform into a model in fsQCA software. For example, a proposition can be the following: *Customers having high quality of personalization, high message quality, and high strongly positive emotions will have high intention to purchase online.* Thus, the model we need to compute is the presence of *quality of personalization, message quality and positive emotions*. Indeed, the model should be seen as one variable, thus we use the option *Compute* from the *Variable* menu. The function *fuzzyand(x,..,)* is used, which takes as input all the variables that are present (Fig. 11). Also, here we can test for the negation of a specific condition if it is required by a proposition. In this case, first we use the *fuzzynot(x)*

Table 5
A summary of thresholds used in fsQCA.

	Threshold	Sources	Examples
Any type of data (including Likert Scales)	95 % - Full set membership 50 % - Intermediate membership 5 % - Full set non-membership	(Fiss, 2011; Ragin, 2008b; Rihoux & Ragin, 2009)	
7-point Likert Scale	6 - Full set membership 4 - Intermediate membership 2 - Full set non-membership	(Ordanini et al., 2014)	(Pappas et al., 2016, 2020)
Any type of data skewed to the right or left (e.g., Likert scale, clickstreams, user performance, physiological data)	80 % - Full set membership 50 % - Intermediate membership 20 % - Full set non-membership		(Pappas, Mikalef et al., 2017); *Likert Scale* (Papamitsiou et al., 2020); *Clickstreams, multimodal data, user performance*
Overall solution consistency	> 0.75 minimum		
Raw consistency	> 0.80 suggested Can be higher Identify natural breaking points in the consistency values Test multiple values of consistency threshold for the same analysis and assess the difference between the results. Low consistency threshold leads to more necessary conditions, reducing type II errors (i.e., false negatives), but increasing type I errors (i.e., false positives), and vice versa.	(Dul, 2016; Fiss, 2011; Mendel & Korjani, 2012; Ragin, 2008b; Rihoux & Ragin, 2009)	
PRI consistency	Close to "Raw consistency" (~0.70) > 0.50 minimum	(Greckhamer et al., 2018)	
Coverage	No specific threshold For Small-N sample size high coverage is expected.	(Greckhamer et al., 2013)	
Sample Size	<50: Small-N >50: Large-N No restriction on how big the sample can be	(Greckhamer et al., 2013; Vatrapu et al., 2016)	
Frequency	1 for Small-N 2 or 3 Large-N It is recommended keeping about 80 % of the cases in the analysis.		<60 (Capatina, Micu, Micu, Bouzaabia, & Bouzaabia, 2018) 500+ (Pappas et al., 2016) 3000+ (Schmitt, Grawe, & Woodside, 2017; Sergis et al., 2018)

for every variable that is negated, which computes the negation (1-x) of a variable (fuzzy set). Then we use the *fyzzyand* function which takes as input all the variables that are present in each configuration and the new variables that occurred as the outcome of the *fuzzynot(x)* function. The *fuzzyand(x,…,)* function returns a minimum of two variables (fuzzy sets).

Finally, the new variable (model) is plotted against the outcome of interest using the XY Plot option from the fsQCA menu *(Graphs – XY Plot)* (Fig. 12). Consistency and coverage values are presented here. The findings show that the proposition is partially supported. Indeed, it does not correspond to a specific solution identified by fsQCA, instead it allows the identification of specific cases, or persons, (who and how many) that will have high or low/medium intentions depending on specific antecedent conditions (if they are high or low/medium). Asymmetric analysis indicates that high scores on the model (i.e., configuration) usually occur for high scores on the outcome condition, making the model useful for the researchers. However, this model does not predict all cases with high scores on the outcome, as other models exist that predict high scores of the same outcome (i.e., the upper left side in the plot). Models with consistency above 0.80 are useful and can serve theory advancement (Woodside, 2017).

6.7. Testing for predictive validity

Testing our solutions (models) for predictive validity is important. Predictive validity shows how well the model predicts the dependent variable in additional samples (Gigerenzer & Brighton, 2009; Gigerenzer & Gaissmaier, 2011; Woodside, 2014). Predictive validity is important because achieving only good model fit does not necessarily mean that the model offers good predictions. We present here how to perform predictive validity testing in fsQCA (Pappas et al., 2016). This process is visualized in Fig. 13.

In order to test for predictive validity, the first step is to randomly divide the sample into a subsample and a holdout sample, and run the same analysis for both samples, as was described in the previous sections. The second step is to run the fsQCA for the subsample, and then the findings obtained should be tested against the holdout sample. Testing for predictive validity including hold-out samples is always possible and doing so substantially increases the added value for both empirical positivistic and interpretative case studies (Woodside, 2010). Table 2 presents the solutions from the subsample.

After obtaining the findings from the subsample, the researcher must use the holdout sample to proceed with predictive validity testing. From the findings of the subsample, each solution that contains the various combinations of present and absent variables, needs modelling as one variable following a similar procedure as described in Section 6.6 (using functions *fuzzyand(x,…,) fuzzynot(x)*. Instead of computing a predesigned proposition we compute every solution from the findings from the testing with the subsample. Finally, the new variable is plotted against the outcome of interest using the holdout sample (Fig. 14). Consistency and coverage values are presented here, which should not contradict the consistency and coverage of the solution. The numbers below the "Plot" button show set-theoretic consistency scores (Ragin, 2018). If one of these two numbers indicates high consistency, the other can be interpreted as a coverage score. In our example, 0.892873 indicates high consistency, while 0.46756 indicates the coverage. These calculations indicate that the data are largely consistent (89 %) with the argument that Model 1 is a subset of intention to purchase and its coverage of intention to purchase is 47 %. That is, Model 1 accounts for 47 % of the sum of the memberships in intention to purchase.

7. When to use fsQCA in preference to variance-based approaches

FsQCA can be complementary to the traditional variance-based approaches and the researchers can decide which method to use in their study, depending on various factors since fsQCA can help overcome some limitations of variance-based approaches. Researchers can perform fsQCA to determine if causal asymmetry exists in their datasets and if their findings are subject to causal equifinality or asymmetry (Fiss, 2011).

This section presents findings from a PLS-SEM analysis on the same constructs and dataset in order to compare them with the fsQCA findings that are presented earlier in the paper. Here, PLS-SEM, CB-SEM, or MRA would lead to similar results as all independent variables are set as predictors of one dependent variable. In detail, Fig. 15 shows the findings from PLS-SEM analysis, which was performed using SmartPLS software, with five out of seven relations being significant. Quality of personalization and benefits of personalization have the strongest influence on intention to purchase, suggesting that they are the most important factors in this model. Also, message quality and weakly positive emotions have a weaker positive effect on intention to purchase, while weakly negative emotions have a negative effect in intention to purchase. Strongly positive and negative emotions do not influence intention to purchase. Finally, the model explains 57 % of the variance of intention to purchase ($R^2 = 0.57$).

Overall, a typical paper could report and explain these findings, connect to previous studies and also discuss some unexpected results such as the non-significant effect of strongly positive emotions. Furthermore, on average the explained variance of 57 % is considered as acceptable and a good result. However, the question remains on how we can explain the rest of the sample (or its variance), as well as on how we can get more details on why we had unexpected results, such as the weak effect of message quality on intention to purchase or the non-significant effect of positive emotions. Different models may explain better our findings, as consumers' emotions when using personalized services may also have either a mediating (Pappas, Kourouthanassis, Giannakos, & Chrissikopoulos, 2014) or moderating (Pappas, Kourouthanassis, Giannakos, & Chrissikopoulos, 2017) role. Nonetheless, the purpose of this analysis is to demonstrate how a variance-based approach differs in practice with the configurational approach. Since these methods can complement each other, the researchers may perform both of them in a paper if that is suitable [e.g., (Fang, Shao, & Wen, 2016; Mikalef & Pateli, 2017)].

Looking the findings from the fsQCA (Table 1) and from the PLS-SEM (Fig. 15), we can make several observations, highlighting how the two methods are complementary, but more importantly how we are able to get more insight into the dataset by employing fsQCA. Table 3 presents a detailed summary of how we can read the findings from the two methods. The interpretation of the findings in fsQCA is an elaborate process due to the very rich information that the analysis provides. This means that researchers can choose on which part of the findings to focus more depending on the context as well as their knowledge on the topic.

In detail, the "PLS-SEM findings" column presents the typical way of reading the findings from such a model. Based on that, the "FsQCA findings" column presents what can be considered as the equivalent finding in configurational analysis. It must be noted that an overlap may occur between the fsQCA findings, which is done here on purpose to highlight the difference with the PLS-SEM findings as well as the rich information that fsQCA provides.

Overall, the findings from PLS-SEM show that quality of personalization, message quality, benefits of personalization, weakly positive emotions, and weakly negative emotions are determinants of intention to purchase. FsQCA reveals similar findings, but more importantly it also identifies conditions that are (1) sufficient or necessary to explain the outcome and (2) insufficient on their own but are necessary parts of solutions that can explain the result (INUS conditions; insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result

The literature on fsQCA is increasing and more researchers employ configurational analysis to get a deeper understanding of the phenomena under examination. However, as the two methods can complement each other, it means that they look at a phenomenon from a different point of view. For more details on best practices on QCA we direct the reader to existing works (Greckhamer et al., 2018; Schneider & Wagemann, 2010; Woodside, 2016b). In Table 4, we summarize some conceptual differences between fsQCA and typical RBMs, related to key assumptions between the two methods, as well as how the analysis is performed and how the findings are presented and interpreted.

Furthermore, Table 5 summarizes some of the frequently used thresholds during the analysis. In fsQCA, the value of 1 defines the full-set membership, 0 defines the full set non-membership and 0.5 the intermediate membership. Because fsQCA uses log-odds, thus not capable of computing memberships that are exactly 0 or 1, the membership thresholds are set as follows; 0.95, 0.50, 0.05. Nonetheless, the researchers can choose different thresholds depending on the background, context and previous knowledge. Finally, to explicate how existing studies may benefit by applying fsQCA, in Appendix D we summarize their main outcomes and suggest possible extensions through fsQCA based on their findings.

8. Conclusion

The choice of an appropriate method of analysis matters and such choices are defined by our research questions and research objectives. Nonetheless, various limitations may influence these choices such as sample size limitations, availability of tools, or sometimes our own knowledge of specific tools and methods. Often, we tend to keep using the same tools and methods of analysis because we have gained a certain amount of expertise along with the level of convenience it brings. While gaining expertise on a tool or topic is much desired, we may avoid using new tools and methods because they require extra effort and resources to get past their learning curve, something that we may often miss as academics. Methodology defines how we study a phenomenon and how we think about it (Bagozzi, 2007). While quantitative and qualitative approaches have their strengths and weaknesses, employing mixed method approaches can offer significantly deeper insight into our datasets, and by extension the phenomena we study. FsQCA is tool that combines aspects of both quantitative and qualitative approaches in one analysis, thus bridging the quantitative-qualitative gap that exists in most fields.

Similar to Gefen, Straub, and Boudreau (2000) seminal study on how to conduct SEM and regression in research, our goal is to offer an easy to follow guide that describes step-by-step how to employ fsQCA in typical studies in information management and marketing. This tutorial can also be followed by studies in other fields in which similar methods are used.

Looking at the literature, while the number of articles employing fsQCA is increasing, it is evident that some venues (journals or conferences) are more frequented than others. This mainly happens because reviewers (and editors) with more expertise are available in these venues, thus making it possible to better understand the method and its implications. This creates a natural cluster with researchers employing fsQCA looking for venues that have already published articles employing the method. Nonetheless, more and more journals and conferences will accept such articles showing that the field is evolving and interest towards less popular methods is increasing. Editors and reviewers can direct authors towards employing fsQCA and taking a look at their data from a different point of view. It must be clear that fsQCA is not a solution to all problems and is not always appropriate. The authors should justify the reasons for employing fsQCA and always follow the recommended thresholds and guidelines, in order to offer meaningful and valid results. FsQCA should not be employing blindly in a mechanical manner. Based on our own personal experience, editors and reviewers will disagree or will not be able to decide about the appropriateness of even the methodological soundness of a manuscript. As with the tutorial on SEM 20 years ago more knowledgeable opinions should be weighted more heavily than those of less understanding (Gefen et al., 2000). Our goal with this tutorial is to help researchers better understand fsQCA, either they act as authors, or as editors and reviewers, and offer a practical guide that can answer many questions for those less familiar with QCA.

CRediT authorship contribution statement

Ilias O. Pappas: Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft. **Arch G. Woodside:** Methodology, Writing - review & editing.

Acknowledgements

The authors would like to thank the Editor of IJIM and the anonymous reviewers for their helpful developmental comments on the work presented in this manuscript. The authors express their appreciation to Katja Bley for her critical reading and thoughtful suggestions during writing the final version of manuscript. We thank Peer Fiss for offering his insight while developing this tutorial article, as well as Michail Giannakos, Panos Kourouthanassis, Patrick Mikalef for helping in earlier phases of this project. Finally, we thank the workshop participants at Swansea University, American College of Greece, and University of Agder for their useful feedback and insightful suggestions. This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Marie Sklodowska-Curie Grant Agreements Nos. 751550.

Appendix A. Construct and scale items

Table A1

Table A1
Scale items with mean, standard deviation and standardized loading.

Construct and scale items	Mean	S.D.	Loading
Quality of Personalization (CA = 0.89)			
How well and efficiently are perceived the offered services, to fulfill customers' needs			
1. Online vendors can provide me with personalized deals/ads tailored to my activity context.	4.6	1.44	0.84
2. Online vendors can provide me with more relevant promotional information tailored to my preferences or personal interests.	4.6	1.36	0.88
3. Online vendors can provide me with the kind of deals/ads that I might like.	4.5	1.32	0.84
Message Quality (CA = 0.93)			
Customer's general perception of the accuracy and completeness of Website information as it relates to products and transactions, when using personalized services.			
1. Personalized services provide correct information about items or services I want to purchase.	4.3	1.30	0.74
2. Overall, I think personalized services provide useful information.	4.5	1.32	0.80
3. Personalized services provide timely information on an item/service.	4.5	1.29	0.76
4. Personalized services provide sufficient information when I try to make an online purchase.	4.3	1.33	0.77
5. I am satisfied with the information that personalized services provide.	4.6	1.40	0.84
6. Overall, the information personalized services provide is of high quality.	4.4	1.43	0.89
7. Personalized services provide timely information on an item/service.	4.3	1.31	0.86
Benefits of Personalization (CA = 0.90)			
Customer's belief about the extent to which he or she will become better off from the online transaction with a certain Website, when using personalized services.			
1. I think the use of personalized services is convenient.	4.9	1.43	0.85
2. I can save money by using personalized services.	4.7	1.60	0.76
3. I can save time by using personalized services.	5.2	1.57	0.86
4. Using personalized services enables me to accomplish a shopping task more quickly than using traditional methods.	5.9	1.57	0.84
5. Using personalized services increases my productivity in shopping (e.g., make purchase decisions or find product information within the shortest time frame).	4.8	1.57	0.74
Intention to Purchase (CA = 0.90)			
Customer's intention to shop online based on personalized services.			
1. In the future I intend to continue shopping online based on personalized services.	4.7	1.47	0.93
2. My general intention to buy online based on personalized services is very high.	4.4	1.54	0.94
3. I will shop online in the future based on personalized services.	4.4	1.42	0.87

Emotions

Measuring customer's emotions, based on valence and control, when using personalized services.

	Mean	SD	Loading		Mean	SD	Loading
Strongly Positive (CA = 0.91)				**Weakly Positive (CA = 0.89)**			
1. Pleasure	3.7	1.73	0.95	1. Contentment	3.7	1.73	0.93
2. Joy	3.5	1.71	0.93	2. Admiration	3.1	1.73	0.76
3. Pride	2.8	1.63	0.74	3. Love	2.4	1.57	0.64
4. Amusement	3.8	1.57	0.75	4. Relief	3.0	1.79	0.75
5. Interest	4.21	1.51	0.68				
Strongly Negative (CA = 0.83)				**Weakly Negative (CA=0.89)**			
1. Anger	2.99	1.71	0.72	1. Disappointment	2.72	1.65	0.68
2. Hate	2.66	1.45	0.85	2. Shame	2.12	1.48	0.87
3. Contempt	3.09	1.83	.65	3. Regret	2.47	1.61	0.80
4. Disgust	2.15	1.39	0.86	4. Guilt	2.04	1.33	0.84
5. Fear	2.79	1.72	0.57	5. Sadness	1.93	1.32	0.77
				6. Compassion	2.12	1.38	0.69

CA = Cronbach's alpha.

Appendix B. Contrarian case analysis

We create quintiles [i.e., dividing the sample into five equal groups] by ranking the cases. An easy way to do this in SPSS in using the "Rank Cases" function of SPSS. In detail, on menu "*Transform*", select "*Rank Cases…*". Next, under "Rank Types…" uncheck "*Rank*" and select the "*Ntiles*" option and set it to 5 (i.e., quintile) (Fig. B1). Then select the variables of interest and press OK to create the Ntiles.

This leads to the creation of new variables in the dataset (Fig. B2). The quintiles can be also computed by using the percentiles within the descriptive statistics function of SPSS; however, the Rank Cases option is simpler

Next, we performed cross-tabulations across the quintiles, using the SPSS "*Crosstabs*" function (Analyze > Descriptive Statistics > Crosstabs), between every independent variable and the dependent variable (Fig. B3). The crosstabs allow us to compute the degree of association between the variables, which suggests a dependence between the two variables and describes main effects between them.

The result for any two variables is a 5 × 5 table that presents all combinations for all of the cases in the sample between the two variables (Fig. B4). The top left and bottom right cases represent the main effects (e.g., degree of association), while the bottom left and top right represent contrarian cases, that are the ones not explained by the main effects. Results for the contrarian case analysis for all variables are presented in Appendix C, as it appears on the original study of Pappas et al. (2016). The findings show the existence of various relationships between the variables, separate from the main effect, supporting the need to perform a configurational analysis.

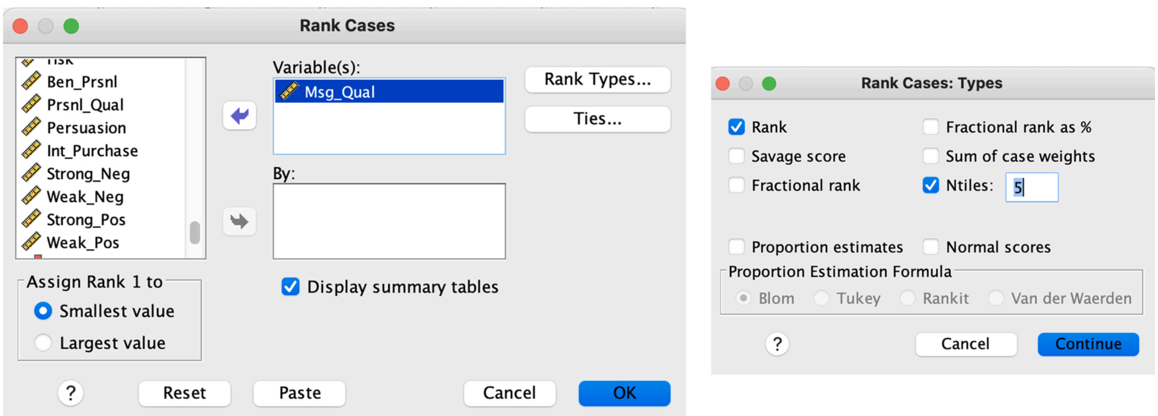


Fig. B1. Create the quintiles using the Rank Cases function in SPSS.

NMsf_Qua	Numeric	3	0	Percentile Group of Msg_Quality	None	None	10	Right	Ordinal	Input
NBenefit	Numeric	3	0	Percentile Group of Benefits_Prsnl	None	None	8	Right	Ordinal	Input
NPrsnl_Q	Numeric	3	0	Percentile Group of Prsnl_Quality	None	None	10	Right	Ordinal	Input
NIntention	Numeric	3	0	Percentile Group of Intention_Purchase	None	None	10	Right	Ordinal	Input
NStr_Neg	Numeric	3	0	Percentile Group of Strong_Negative	None	None	10	Right	Ordinal	Input
NWeak_Neg	Numeric	3	0	Percentile Group of Weak_Negative	None	None	10	Right	Ordinal	Input
NStr_Pos	Numeric	3	0	Percentile Group of Strong_Positive	None	None	8	Right	Ordinal	Input
NWeak_Pos	Numeric	3	0	Percentile Group of Weak_Positive	None	None	10	Right	Ordinal	Input

Fig. B2. Quintiles as new variables in the dataset.

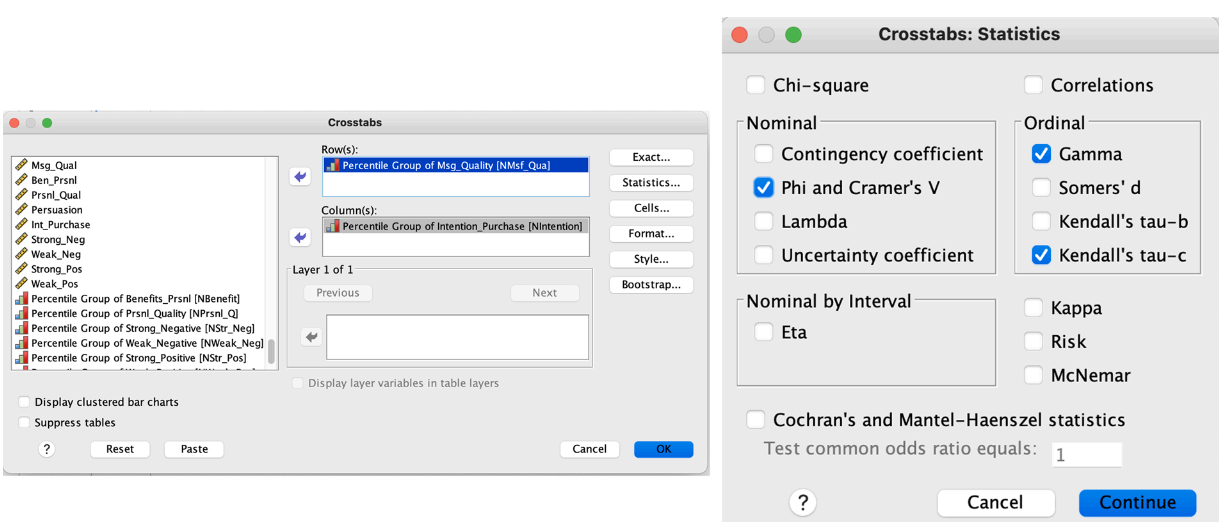


Fig. B3. Performing crosstabs in SPSS to find the main effects and contrarian cases.

Percentile Group of Msg_Quality * Percentile Group of Intention_Purchase Crosstabulation

Count

		Percentile Group of Intention_Purchase					
		1	2	3	4	5	Total
Percentile Group of Msg_Quality	1	62	29	10	5	3	109
	2	32	38	25	14	11	120
	3	25	20	28	26	15	114
	4	9	13	34	52	26	134
	5	5	8	9	43	40	105
Total		133	108	106	140	95	582

Symmetric Measures

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Nominal by Nominal	Phi	.626			.000
	Cramer's V	.313			.000
Ordinal by Ordinal	Kendall's tau-c	.464	.027	17.184	.000
	Gamma	.569	.032	17.184	.000
N of Valid Cases		582			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Fig. B4. Cross-tabulation and degree of association for message quality and intention to purchase.

Appendix C. Results of contrarian case analysis

Table C1

Table C1
Results from the contrarian case analysis.

		Intention to Purchase							Intention to Purchase				
		1	2	3	4	5			1	2	3	4	5
Q.P. (phi^2 =.43, p < .001)	1	*65 (11.2%)*	*21 (3.6%)*	6 (1.0%)	**8 (1.4%)**	**4 (0.7%)**	M.Q. (phi^2 =.50, p < .001)	1	*62 (10.7%)*	*29 (5%)*	10 (1.7%)	**5 (0.9%)**	**3 (0.5%)**
	2	*33 (5.7%)*	*33 (5.7%)*	33 (5.7%)	**12 (2.1%)**	**7 (1.2%)**		2	*32 (5.5%)*	*38 (6.5%)*	25 (4.3%)	**14 (2.4%)**	**11 (1.9%)**
	3	14 (2.4%)	25 (4.3%)	22 (3.8%)	22 (3.8%)	8 (1.4%)		3	25 (4.3%)	20 (3.4%)	28 (4.8%)	26 (4.5%)	15 (2.6%)
	4	**18 (3.1%)**	**24 (4.1%)**	33 (5.7%)	*65 (11.2%)*	*34 (5.8%)*		4	**9 (1.5%)**	**13 (2.2%)**	34 (5.8%)	*52 (8.9%)*	*26 (4.5%)*
	5	**3 (0.5%)**	**5 (0.9%)**	12 (2.1%)	*33 (5.7%)*	*42 (7.2%)*		5	**5 (0.9%)**	**8 (1.4%)**	9 (1.5%)	*43 (7.4%)*	*40 (6.9%)*
B.P. (phi^2 =.39, p < .001)	1	*72 (12.4%)*	*25 (4.3%)*	12 (2.1%)	**4 (0.7%)**	**0 (0.0%)**	S.P. Emotions (phi^2 =.16, p < .001)	1	*55 (9.5%)*	*23 (4%)*	17 (2.9%)	**26 (4.5%)**	**5 (0.9%)**
	2	*33 (5.7%)*	*39 (6.7%)*	23 (4.0%)	**18 (3.1%)**	**4 (0.7%)**		2	*25 (4.3%)*	*24 (4.1%)*	16 (2.7%)	**26 (4.5%)**	**15 (2.6%)**
	3	17 (2.9%)	19 (3.3%)	35 (6.0%)	37 (6.4%)	17 (2.9%)		3	30 (5.2%)	31 (5.3%)	26 (4.5%)	32 (5.5%)	15 (2.6%)
	4	**5 (0.9%)**	**16 (2.7%)**	28 (4.8%)	*42 (7.2%)*	*23 (4.0%)*		4	**17 (2.9%)**	**17 (2.9%)**	23 (4%)	*26 (4.5%)*	*20 (3.4%)*
	5	**6 (1.0%)**	**9 (1.5%)**	8 (1.4%)	*39 (6.7%)*	*51 (8.8%)*		5	**6 (1.0%)**	**13 (2.2%)**	24 (4.1%)	*30 (5.2%)*	*40 (6.9%)*
W.P. Emotions (phi^2 =.14, p < .001)	1	*43 (7.4%)*	*18 (3.1%)*	16 (2.7%)	**20 (3.4%)**	**5 (0.9%)**	S.N. Emotions (phi^2 =.10, p < .001)	1	**22 (3.8%)**	**14 (2.4%)**	13 (2.2%)	*41 (7%)*	*35 (6.0%)*
	2	*28 (4.8%)*	*29 (5%)*	18 (3.1%)	**31 (5.3%)**	**17 (2.9%)**		2	**18 (3.1%)**	**16 (2.7%)**	21 (3.6%)	*37 (6.4%)*	*20 (3.4%)*
	3	26 (4.5%)	30 5.2%)	26 (4.5%)	35 (6%)	13 (2.2%)		3	28 (4.8%)	28 (4.8%)	22 (3.8%)	24 (4.1%)	22 (3.8%)
	4	**25 (4.3%)**	**23 (4%)**	19 (3.3%)	*22 (3.8%)*	*19 (3.3%)*		4	*33 (5.7%)*	*22 (3.8%)*	25 (4.3%)	**17 (2.9%)**	**7 (1.2%)**
	5	**11 (1.9%)**	**8 (1.4%)**	27 (4.6%)	*32 (5.5%)*	*41 (7%)*		5	*32 (5.5%)*	*28 (4.8%)*	25 (4.3%)	**21 (3.6%)**	**11 (1.9%)**
W.N. Emotions (phi^2 =.05, p < .05)	1	**24 (4.1%)**	**20 (3.4%)**	19 (3.3%)	*37 (6.4%)*	*28 (4.8%)*							
	2	**19 (3.3%)**	**10 (1.7%)**	12 (2.1%)	*33 (5.7%)*	*19 (3.3%)*							
	3	33 (5.7%)	25 (4.3%)	26 (4.5%)	24 (4.1%)	20 (3.4%)							
	4	*26 (4.5%)*	*31 (5.3%)*	26 (4.5%)	**21 (3.62%)**	**16 (2.7%)**							
	5	*31 (5.3%)*	*22 (3.8%)*	23 (4%)	**25 (4.3%)**	**12 (2.1%)**							

Cases in **bold** represent contrarian cases. Cases in *italics* represent main effect.

The sets of contrarian cases are counter to the main effect size (phi^2 range from .05 to .50).

QP; Quality of Personalization, BP; Benefits of Personalization, MQ; Message Quality, SP; Strongly Positive Emotions, WP; Weakly Positive Emotions, SN; Strongly Negative Emotions, WN; Weakly Negative

Appendix D. Possible extension of recent studies with fsQCA

Table D1

Table D1
Examples of recent studies that perform variance-based analysis (e.g., MRA, SEM) along with suggestions for possible extensions with fsQCA.

Source	Topic Description	Main findings	Possible extension with fsQCA
Aladwani and Dwivedi (2018)	Citizen-government interaction via social media. The article proposes and tests a new model in a quest for a SocioCitizenry theory.	Anticipated quality, configured trust, and approved adaptation as important factors in citizen-government interaction. Anticipated governmental social media quality influences configured trust, which in turn influences the extent of approved adaptation.	Examine if the dimensions of anticipated quality and configured trust, are indispensable factors to achieve approved adaptation. Identify which of these dimensions are sufficient to explain high approved adaptation.
Veeramootoo, Nunkoo, and Dwivedi (2018)	Continuance usage intention of e-filing in e-government	System quality, user satisfaction and habit positively influenced by intention to continue using e-filing. Satisfaction is the most important factor having the stronger impact on the outcome. Information quality, service quality and perceived risk do not significantly predict continuance usage intention of e-filing.	Examine if satisfaction is a necessary factor for high continuance usage intention of e-filing or if it needs to be combined with other factors. Identify if combinations of system quality, satisfaction, and habit are sufficient to explain the outcome. Identify how information quality, service quality, and perceived risk play a role for a smaller part of the sample or if they help to formulate solutions that are different from the main effects.
Hossain, Dwivedi, Chan, Standing, and Olanrewaju (2018)	Sharing Political Content in Online Social Media	The factors representing both planned (i.e., perceived social recognition and altruistic motivation) and unplanned behavior (i.e., extroversion and impulsiveness) affect people's political content sharing behavior. Collective opinion moderates planned behavior, but not unplanned behavior.	Examine how the absence of planned behavior and absence of unplanned behavior can explain political content sharing behavior as well as its negation (i.e., not sharing). Complement the role of collective opinion as moderator by identifying the combinations in which it is necessary to be present or absent for people to share political content.
Kamboj, Sarmah, Gupta, and Dwivedi (2018)	Customer participation in social media brand communities applying the S-O-R framework.	SNSs participation motivations significantly affect customer participation, which in turn positively influences brand trust, brand loyalty and consequently resulted in branding co-creation. Gender influences brand trust and brand co-creation, while age influences brand loyalty Brand trust mediates the relationship between customer participation and brand loyalty in social media brand communities.	Examine how participation, brand trust, brand loyalty can combine to explain brand co-creation for different customers depending on their gender and age. Examine how brand loyalty in social media brand communities' changes when brand trust is present and when it is absent.
Dwivedi, Shareef, Mukerji, Rana, and Kapoor (2018)	Exploratory study in involvement in emergency supply chain for disaster management	Administrative conflict, political biasness and professional growth have significant effects on negative attitude. Impact of insecurity and corruption is non-significant on attitude. Insecurity and Corruption constructs were removed from the model to improve the model the fit. Negative attitude has a very strong effect on behavioral intentions.	Test the same antecedents for both positive and negated attitude and compare. Conditions that explain the presence of the outcome are not necessarily mirror opposites of those explaining the absence of the outcome. Include insecurity and corruption as they may play a role in a small part of the sample (not explain by the main model). Examine if negative attitude is a necessary factor for behavioral intentions.
Dwivedi et al. (2017)	Propose and test the unified model of electronic government adoption (UMEGA)	UMEGA outperforms all other models for e-government Government context should be taken into account. UMEGA is simpler to use and has a better explanatory power than the UTAUT.	Identify how the antecedents form different solutions that explain attitude and intention separately and compare. Examine if attitude is a necessary factor for high intention, and if not, what other combinations are sufficient to explain intention that do not include attitude. Since the explanatory power is so large complement by additional solutions for the small part of the sample that is not explained by this model.
Dwivedi et al. (2019)	Propose a revised alternative theoretical model for explaining the acceptance and use of IS and IT innovations using a combination of meta-analysis and structural equation modelling (MASEM) techniques.	Attitude was central to behavioral intentions and usage behaviors, partially mediated the effects of exogenous constructs on behavioral intentions, and had a direct influence on usage behaviors.	Identify how the antecedents form different solutions that are sufficient to explain attitude and intention separately and compare. Examine if attitude is a necessary factor for high intention, and if not what other combinations are sufficient to explain intention that do not include attitude.

(*continued on next page*)

Table D1 (*continued*)

Source	Topic Description	Main findings	Possible extension with fsQCA
Rana et al. (2017)	Adoption of emerging electronic government (eGov) applications in India	Performance expectancy, effort expectancy, social influence, facilitating conditions, anxiety are important predictors of attitude and behavioral intention.	Test specific propositions coming from MASEM and identify specific cases in the sample. Since the model performs better than other acceptance models, identify if any of the antecedents is a necessary condition for the outcome to occur.
		The model performs better compare to previous technology acceptance models.	Identify combinations that explain a different (smaller) part of the sample that is not represented in the main model.
Shareef, Dwivedi, Kumar, and Kumar (2017)	Examine the effects of promotional marketing through SMS.	It revealed that consumer segmentation and target marketing is the most effective way to communicate with consumers through promotional marketing conducted by the mobile phone SMS.	Examine if consumer segmentation and target marketing are necessary in order to effectively communicate with consumers through promotional marketing conducted by the mobile phone SMS.
		It also suggested that this promotional marketing is valuable only for highly reputable vendors/ retailers.	Identify different combinations of consumer segmentation and target marketing that are sufficient to explain the outcome.
Alalwan et al. (2017)	The study examines the factors influencing behavioral intentions and adoption of Mobile banking by Jordanian bank customers.	All predictors (Performance Expectancy, Effort Expectancy, Trust, Hedonic Motivation, Price Value) of behavioral intention to adopt mobile banking were found significant except for social influence.	Examine if all predictors (except SI) of behavioral intention are necessary to be present in order to explain adoption of mobile banking.
		The variance explained by the model in behavioral intention was found as 65 %.	Identify combinations that explain other part of the sample, not described in the main model that explains 65 % of the variance of behavioral intention.
Shareef, Dwivedi, Kumar, and Kumar (2016)	This study addresses whether a service delivery channel, based on mobile phones and provided through a short messaging service (SMS) can be included in public administration to meet citizen requirements, regarding their perceptions of high value and effectiveness.	Citizens are quite satisfied with SMSs used as a service delivery channel by the public service domain.	Identify which factors are necessary or sufficient to develop citizen perceptions of high value and effectiveness of this new service delivery channel.
	Also, the study examines if culture plays a role on developing a positive attitude toward this service delivery channel.	If public service providers can effectively segment the market based on time, location, and requirements, and can deliver the preferred message to concerned users with relevant and information that is easy to access and process, citizens will regard this service delivery channel as effective and satisfactory, and as competent as its private counterparts.	Produce solutions that profile citizens based on their country and identify which factors are necessary or sufficient for high positive attitudes towards this service for the different cultures. The relatively small sample per country allows to go back to the cases and explain behavior using contextual information
Rana, Dwivedi, Williams, and Weerakkody (2016)	Adoption of online public grievance redressal system in India	All are verified.	Identify how the antecedents form different solutions that explain attitude and intention separately and compare.
		The model performs better compared to previous technology acceptance models.	Examine if attitude is a necessary factor for high intention, and if not what other combinations are sufficient to explain intention that do not include attitude.
Slade, Dwivedi, Piercy, and Williams (2015)	Adoption of Remote Mobile Payments (MP) in the UK. Extending the Unified Theory of Acceptance and Use of Technology (UTAUT), with more consumer-related constructs (Innovativeness, Risk, and Trust) for non-users of these services.	Performance expectancy, social influence, innovativeness, and perceived risk influence behavioral intentions.	Identify which of the proposed antecedents are sufficient or necessary for high behavioral intentions.
		Effort expectancy and trust do not influence behavioral intentions.	Make propositions based on the SEM findings and test them to identify for how many cases in the sample (and which ones) they hold true.
		Knowledge of MP moderates the effects of antecedents of behavioral intention	Segment consumers based on their knowledge of MP, identify who they are and use knowledge that is not included in the model to explain their behavior. The latter could lead to qualitative studies with specific consumers (e.g., via interviews).

References

Aladwani, A. M., & Dwivedi, Y. K. (2018). Towards a theory of SocioCitizenry: Quality anticipation, trust configuration, and approved adaptation of governmental social media. *International Journal of Information Management, 43*, 261–272.

Alalwan, A. A., Dwivedi, Y. K., & Rana, N. P. (2017). Factors influencing adoption of mobile banking by Jordanian bank customers: Extending UTAUT2 with trust. *International Journal of Information Management, 37*(3), 99–110.

Bagozzi, R. P. (2007). The legacy of the technology acceptance model and a proposal for a paradigm shift. *Journal of the Association for Information Systems, 8*(4), 3.

Basurto, X., & Speer, J. (2012). Structuring the calibration of qualitative data as sets for qualitative comparative analysis (QCA). *Field Methods, 24*(2), 155–174.

Berg-Schlosser, D., & De Meur, G. (2009). Comparative research design: Case and variable selection. In B. Rihoux, & C. C. Ragin (Eds.), *Configurational comparative methods. Qualitative comparative analysis (QCA) and related techniques* (pp. 19–32). Thousand Oaks and London: Sage.

Cairns, J., Wistow, J., & Bambra, C. (2017). Making the case for qualitative comparative analysis in geographical research: A case study of health resilience. *Area, 49*(3), 369–376.

Capatina, A., Micu, A., Micu, A. E., Bouzaabia, R., & Bouzaabia, O. (2018). Country-based comparison of accommodation brands in social media: An fsQCA approach. *Journal of Business Research, 89*, 235–242.

Cooper, B., & Glaesser, J. (2011). Using case-based approaches to analyse large datasets: A comparison of Ragin's fsQCA and fuzzy cluster analysis. *International Journal of Social Research Methodology, 14*(1), 31–48.

Cronqvist, L. (2004). *Presentation of TOSMANA. Adding multi-value variables and visual aids to QCA. COMPASSS: WP,20*.

Cronqvist, L., & Berg-Schlosser, D. (2009). Multi-value QCA (mvQCA). *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques, 51*, 69–86.

Delery, J. E., & Doty, D. H. (1996). Modes of theorizing in strategic human resource management: Tests of universalistic, contingency, and configurational performance predictions. *The Academy of Management Journal, 39*(4), 802–835.

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation, 14*(1), 20.

Dul, J. (2016). Identifying single necessary conditions with NCA and fsQCA. *Journal of Business Research, 69*(4), 1516–1523.

Dwivedi, Y. K., Rana, N. P., Janssen, M., Lal, B., Williams, M. D., & Clement, M. (2017). An empirical validation of a unified model of electronic government adoption (UMEGA). *Government Information Quarterly, 34*(2), 211–230.

Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., & Williams, M. D. (2019). Re-examining the unified theory of acceptance and use of technology (UTAUT): Towards a revised theoretical model. *Information Systems Frontiers, 21*(3), 719–734.

Dwivedi, Y. K., Shareef, M. A., Mukerji, B., Rana, N. P., & Kapoor, K. K. (2018). Involvement in emergency supply chain for disaster management: A cognitive dissonance perspective. *International Journal of Production Research, 56*(21), 6758–6773.

El Sawy, O. A., Malhotra, A., Park, Y., & Pavlou, P. A. (2010). Research commentary-seeking the configurations of digital ecodynamics: It takes three to tango. *Information Systems Research, 21*(4), 835–848.

Fang, J., Shao, Y., & Wen, C. (2016). Transactional quality, relational quality, and consumer e-loyalty: Evidence from SEM and fsQCA. *International Journal of Information Management, 36*(6), 1205–1217.

Fedorowicz, J., Sawyer, S., & Tomasino, A. (2018). Governance configurations for inter-organizational coordination: A study of public safety networks. *Journal of Information Technology, 33*(4), 326–344.

Fiss, P. C. (2007). A set-theoretic approach to organizational configurations. *The Academy of Management Review, 32*(4), 1180–1198.

Fiss, P. C. (2011). Building better causal theories: A fuzzy set approach to typologies in organization research. *The Academy of Management Journal, 54*(2), 393–420.

Fitzsimons, G. J. (2008). Death to dichotomizing. *The Journal of Consumer Research, 35*(1), 5–8.

Gefen, D., Straub, D., & Boudreau, M.-C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems, 4*(1), 7.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review, 98*(2), 254.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1*(1), 107–143.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451–482.

Greckhamer, T., Furnari, S., Fiss, P. C., & Aguilera, R. V. (2018). Studying configurations with qualitative comparative analysis: Best practices in strategy and organization research. *Strategic Organization, 18*(3), 482–495.

Greckhamer, T., Misangyi, V. F., & Fiss, P. C. (2013). The two QCAs: From a small-N to a large-N set theoretic approach. In P. C. Fiss, B. Cambré, & A. Marx (Eds.), *Configurational theory and methods in organizational research* (pp. 49–75). Emerald Group Publishing Limited.

Henik, E. (2015). Understanding whistle-blowing: A set-theoretic approach. *Journal of Business Research, 68*(2), 442–450.

Hossain, M. A., Dwivedi, Y. K., Chan, C., Standing, C., & Olanrewaju, A.-S. (2018). Sharing political content in online social media: A planned and unplanned behaviour approach. *Information Systems Frontiers, 20*(3), 485–501.

Kamboj, S., Sarmah, B., Gupta, S., & Dwivedi, Y. (2018). Examining branding co-creation in brand communities on social media: Applying the paradigm of Stimulus-Organism-Response. *International Journal of Information Management, 39*, 169–185.

Kane, H., Lewis, M. A., Williams, P. A., & Kahwati, L. C. (2014). Using qualitative comparative analysis to understand and quantify translation and implementation. *Translational Behavioral Medicine, 4*(2), 201–208.

Liu, Y., Mezei, J., Kostakos, V., & Li, H. (2017). Applying configurational analysis to IS behavioural research: A methodological alternative for modelling combinatorial complexities. *Information Systems Journal, 27*(1), 59–89.

Longest, K. C., & Vaisey, S. (2008). Fuzzy: A program for performing qualitative comparative analyses (QCA) in Stata. *The Stata Journal, 8*(1), 79–104.

Mackie, J. I. (1965). Causes and conditions. *American Philosophical Quarterly, 2*(4), 245–264.

Mendel, J. M., & Korjani, M. M. (2012). Charles Ragin's fuzzy set qualitative comparative analysis (fsQCA) used for linguistic summarizations. *Information Sciences, 202*, 1–23.

Mikalef, P., & Pateli, A. (2017). Information technology-enabled dynamic capabilities and their indirect effect on competitive performance: Findings from PLS-SEM and fsQCA. *Journal of Business Research, 70*, 1–16.

Miller, D. (2018). Challenging trends in configuration research: Where are the configurations? *Strategic Organization, 16*(4), 453–469.

Misangyi, V. F., Greckhamer, T., Furnari, S., Fiss, P. C., Crilly, D., & Aguilera, R. (2017). Embracing causal complexity: The emergence of a neo-configurational perspective. *Journal of Management, 43*(1), 255–282.

Nagy, G., Megehee, C. M., Woodside, A. G., Laukkanen, T., Hirvonen, S., & Reijonen, H. (2017). Achieving requisite variety in modeling firms' strategy heterogeneities: Explaining paradoxical firm-market performances. *Industrial Marketing Management, 65*, 100–128.

Nistor, N., Stanciu, D., Lerche, T., & Kiel, E. (2019). "I am fine with any technology, as long as it doesn't make trouble, so that I can concentrate on my study": A case study of university students' attitude strength related to educational technology acceptance. *British Journal of Educational Technology, 50*(5), 2557–2571.

Ordanini, A., Parasuraman, A., & Rubera, G. (2014). When the recipe is more important than the ingredients a Qualitative Comparative Analysis (QCA) of service innovation configurations. *Journal of Service Research, 17*(2), 134–149. https://doi.org/10.1177/1094670513513337.

Papamitsiou, Z., Economides, A. A., Pappas, I. O., & Giannakos, M. N. (2018). Explaining learning performance using response-time, self-regulation and satisfaction from content: An fsQCA approach. *Paper Presented at the Proceedings of the 8th International Conference on Learning Analytics and Knowledge*.

Papamitsiou, Z., Pappas, I. O., Sharma, K., & Giannakos, M. (2020). Utilizing multimodal data through fsQCA to explain engagement in adaptive learning. *IEEE Transactions on Learning Technologies, 13*(4), 689–703. https://doi.org/10.1109/TLT.2020.3020499.

Pappas, I. O. (2018). User experience in personalized online shopping: A fuzzy-set analysis. *European Journal of Marketing, 52*(7/8), 1679–1703. https://doi.org/10.1108/EJM-10-2017-0707.

Pappas, I. O., Kourouthanassis, P. E., Giannakos, M. N., & Chrissikopoulos, V. (2014). Shiny happy people buying: The role of emotions on personalized e-shopping. *Electronic Markets, 24*(3), 193–206.

Pappas, I. O., Kourouthanassis, P. E., Giannakos, M. N., & Chrissikopoulos, V. (2016). Explaining online shopping behavior with fsQCA: The role of cognitive and affective perceptions. *Journal of Business Research, 69*(2), 794–803.

Pappas, I. O., Papavlasopoulou, S., Mikalef, P., & Giannakos, M. N. (2020). Identifying the combinations of motivations and emotions for creating satisfied users in SNSs: An fsQCA approach. *International Journal of Information Management, 53*, Article 102128.

Pappas, I. O., Giannakos, M. N., Jaccheri, L., & Sampson, D. G. (2017). Assessing student behavior in computer science education with an fsQCA approach: The role of gains and barriers. *ACM Transactions on Computing Education (TOCE), 17*(2). Article No. 10.

Pappas, I. O., Kourouthanassis, P. E., Giannakos, M. N., & Chrissikopoulos, V. (2017). Sense and sensibility in personalized e-commerce: How emotions rebalance the purchase intentions of persuaded customers. *Psychology & Marketing, 34*(10), 972–986.

Pappas, I. O., Mikalef, P., Giannakos, M. N., & Pavlou, P. A. (2017). Value co-creation and trust in social commerce: An fsQCA approach. *Paper Presented at the 25th European Conference on Information Systems (ECIS)*.

Park, Y., & Mithas, S. (2020). Organized complexity of digital business strategy: A configurational perspective. *MIS Quarterly, 44*(1).

Park, Y., Fiss, P., & El Sawy, O. A. (2020). Theorizing the multiplicity of digital phenomena: The ecology of configurations, causal recipes, and guidelines for applying QCA. *MIS Quarterly, 44*(4), 1493–1520.

Plewa, C., Ho, J., Conduit, J., & Karpen, I. O. (2016). Reputation in higher education: A fuzzy set analysis of resource configurations. *Journal of Business Research, 69*(8), 3087–3095.

Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Univ of California Press.

Ragin, C. C. (2000). *Fuzzy-set social science*. University of Chicago Press.

Ragin, C. C. (2018). *User's guide to Fuzzy-Set/Qualitative comparative analysis 3.0*. Irvine, California: Department of Sociology, University of California.

Ragin, C. C. (2008a). Measurement versus calibration: A set-theoretic approach. In J. M. Box-Steffensmeier, H. E. Brady, & D. Collier (Eds.), *The Oxford handbook of political methodology* (pp. 174–198).

Ragin, C. C. (2008b). *Redesigning social inquiry: Fuzzy sets and beyond* (Vol. 240). Wiley Online Library.

Ragin, C. C., & Davey, S. (2016). *fs/QCA [Computer Programme], version 3.0*. Irvine, CA: University of California.

Rana, N. P., Dwivedi, Y. K., Lal, B., Williams, M. D., & Clement, M. (2017). Citizens' adoption of an electronic government system: Towards a unified view. *Information Systems Frontiers, 19*(3), 549–568.

Rana, N. P., Dwivedi, Y. K., Williams, M. D., & Weerakkody, V. (2016). Adoption of online public grievance redressal system in India: Toward developing a unified view. *Computers in Human Behavior, 59*, 265–282.

Rihoux, B., & Ragin, C. C. (2009). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (Vol. 51). Thousand Oaks, CA: Sage Publications.

Saridakis, C., Angelidou, S., & Woodside, A. G. (2020). What type of CSR engagement suits my firm best? Evidence from an abductively-derived typology. *Journal of Business Research, 108*, 174–187.

Schmitt, A. K., Grawe, A., & Woodside, A. G. (2017). Illustrating the power of fsQCA in explaining paradoxical consumer environmental orientations. *Psychology & Marketing, 34*(3), 323–334.

Schneider, C. Q., & Wagemann, C. (2010). Standards of good practice in qualitative comparative analysis (QCA) and fuzzy-sets. *Comparative Sociology, 9*(3), 397–418.

Sergis, S., Sampson, D. G., & Giannakos, M. N. (2018). Supporting school leadership decision making with holistic school analytics: Bridging the qualitative-quantitative divide using fuzzy-set qualitative comparative analysis. *Computers in Human Behavior, 89*, 355–366.

Shareef, M. A., Dwivedi, Y. K., Kumar, V., & Kumar, U. (2016). Reformation of public service to meet citizens' needs as customers: Evaluating SMS as an alternative service delivery channel. *Computers in Human Behavior, 61*, 255–270.

Shareef, M. A., Dwivedi, Y. K., Kumar, V., & Kumar, U. (2017). Content design of advertisement for consumer exposure: Mobile marketing through short messaging service. *International Journal of Information Management, 37*(4), 257–268.

Slade, E. L., Dwivedi, Y. K., Piercy, N. C., & Williams, M. D. (2015). Modeling consumers' adoption intentions of remote mobile payments in the United Kingdom: Extending UTAUT with innovativeness, risk, and trust. *Psychology & Marketing, 32*(8), 860–873.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Sage.

Thiem, A. (2013). Clearly crisp, and not fuzzy: A reassessment of the (putative) pitfalls of multi-value QCA. *Field Methods, 25*(2), 197–207.

Thiem, A., & Dusa, A. (2013). QCA: A package for qualitative comparative analysis. *The R Journal, 5*(1), 87–97.
Urry, J. (2005). The complexity turn. *Theory, Culture & Society, 22*(5), 1.
Vatrapu, R., Mukkamala, R. R., Hussain, A., & Flesch, B. (2016). Social set analysis: A set theoretical approach to big data analytics. *IEEE Access, 4*, 2542–2571.
Veeramootoo, N., Nunkoo, R., & Dwivedi, Y. K. (2018). What determines success of an e-government service? Validation of an integrative model of e-filing continuance usage. *Government Information Quarterly, 35*(2), 161–174.
Vink, M. P., & Van Vliet, O. (2009). Not quite crisp, not yet fuzzy? Assessing the potentials and pitfalls of multi-value QCA. *Field Methods, 21*(3), 265–289.
Vink, M. P., & Vliet, O.v. (2013). Potentials and pitfalls of multi-value QCA: Response to Thiem. *Field Methods, 25*(2), 208–213.
Vis, B. (2012). The comparative advantages of fsQCA and regression analysis for moderately large-N analyses. *Sociological Methods & Research, 41*(1), 168–198.
Von Bertalanffy, L. (1968). *General system theory: Foundations, development, applications*. New York: Braziller.
Woodside, A. G. (2010). Bridging the chasm between survey and case study research: Research methods for achieving generalization, accuracy, and complexity. *Industrial Marketing Management, 39*(1), 64–75.
Woodside, A. G. (2013). Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory. *Journal of Business Research, 66*(4), 463–472.
Woodside, A. G. (2014). Embrace• perform• model: Complexity theory, contrarian case analysis, and multiple realities. *Journal of Business Research, 67*(12), 2495–2503.
Woodside, A. G. (2017). *The complexity turn: Cultural, management, and marketing applications*. Berlin: Springer.
Woodside, A. G. (2019). Accurate case-outcome modeling in economics, psychology, and marketing. *Psychology & Marketing, 36*(11), 1046–1061.
Woodside, A. G. (2016a). *Bad to good: Achieving high quality and impact in your research.* Emerald Group Publishing.
Woodside, A. G. (2016b). *The good practices manifesto: Overcoming bad practices pervasive in current research in business*. Elsevier.