Textiles, Fashion Merchandising and Design Faculty Publications

Textiles, Fashion Merchandising and Design

Measuring Perceived Corporate Hypocrisy: Scale Development in the Context of U.S. Retail Employees

Saheli Goswami
University of Rhode Island, sgoswami@uri.edu

Jung Ha-Brookshire

Wes Bonifay

Measuring Perceived Corporate Hypocrisy: Scale Development in the Context of U.S. Retail Employees

Saheli Goswami [1],*, Jung Ha-Brookshire [2] and Wes Bonifay [3]

[1] Department of Textiles, Fashion Merchandising and Design, College of Business, University of Rhode Island, 55 Lower College Road, Kingston, RI 02881, USA

[2] Textile and Apparel Management, College of Human Environmental Sciences, University of Missouri, 137 Stanley Hall, Columbia, MO 65211, USA; habrookshirej@missouri.edu

[3] Educational, School, & Counseling Psychology, University of Missouri, 5C Hill Hall, Columbia, MO 65211, USA; bonifayw@missouri.edu

* Correspondence: sgoswami@uri.edu; Tel.: +1-401-874-9294

Abstract: Despite an increasing amount of research on perceived corporate hypocrisy (PCH), limited research has investigated PCH among employees. Particularly, the literature lacked a valid instrument for estimating employees' PCH, even though employees experience severe consequences for PCH. To address this gap, a scale was developed to measure employees' PCH, using a three-stage Item Response Theory modeling approach. After a series of qualitative studies and six quantitative scale-development iterations, PCH was found to be a unidimensional construct represented by the perceived lack of morality, perceived control breach, double standards, and a value-behavior gap. Further, the nine-item PCH scale was confirmed to be reliable, valid, and unbiased for different demographic groups. The scale makes theoretical contributions by being one of the few attempts to objectively measure employees' hypocrisy judgements and incorporating corporations' double standard and perceived commitment to morality as defining features of employees' PCH. Assessment of employees' PCH can help in expanding the hypocrisy literature beyond consumers' perceptions. Finally, the scale enables corporations to measure employees' PCH and get an in-depth understanding of the issues of concern as work and organizational phenomena. By implementing proper management strategies, corporations can potentially avoid PCH, create more favorable perceptions among employees, and improve their reputations.

Keywords: perceived corporate hypocrisy; scale development; retail; item response theory; employees; behavioral concerns

1. Introduction

Corporations can easily be perceived as hypocritical if they claim to be something that they are not [1]. Particularly in relation to corporate social responsibilities (CSR), businesses are often found engaged in unethical business practices, such as greenwashing, claiming to make social cause-related investments, etc. [2,3]. CSR has also been criticized as "hypocritical window-dressing" [4] with the argument that most corporations use CSR to attract customers and more business [5,6] and has continued to draw scholarly attention [1,7,8]. Corporations often become the subject of media scrutiny not only for failing to behave ethically concerning their consumers, the well-being of people, human rights, or the environment [9], but also for unethical practices, such as false promotions, bait marketing, inaccurate claims, counterfeit products, mislabeling, and privacy infringement for marketing. Such inconsistencies have been perceived or observed across different industry sectors, such as fashion,

food, general merchandise, and others [10], allowing great potential for corporate hypocrisy to be perceived by both consumers and employees.

The role of corporate hypocrisy in corporate sustainability achievements has been highlighted by previous research [11]. Studies have indicated that when corporate wrongdoing or inconsistencies come to the forefront, such word and action misalignment often generate corporate hypocrisy among observants and can lead to negative reactions towards the corporations [1,11]. The issue of PCH has mostly been studied relative to corporate social responsibilities (CSR), as consumers' questionable perceptions of CSR can have negative impact on their brand loyalty, reputation, and equity [12]. Stakeholders cannot be expected to support corporations with hypocritical CSR engagements, i.e., corporations utilizing societal issues for their ulterior motives. Interestingly, while some studies have reported that, through CSR, corporations can gain positive consumer attitudes, corporate image and reputations [13], other studies have reported of the same CSR and sustainability as being commonly used as marketing gimmicks and hence being perceived as the hypocritical window dressing [5]. Such greenwashing may not only generate emotions of betrayal among the consumers, but can also negatively influence their attitudes towards the corporations, with potential loss of trust and increase of skepticism [14]. Applied to employees, while their positive perceptions of employers' CSR can positively impact work engagement [15], their hypocrisy perceptions induced by contradictory CSR can damage corporate sustainability strategies [16]. While it is important for corporations to have clear sustainability goals and structures with no gap between corporate goals and individuals' behavior toward sustainability, perceptions of corporate hypocrisy can challenge corporate sustainability performance in turn, through negative attitudes toward the corporation [11]. Particularly at a time when individuals have quick access to information platforms and social media, corporate wrongdoing and unethical behavior are more easily recognized than ever among stakeholders.

Corporate hypocrisy has been researched to improve corporate image and/or reputation while typically corresponding with corporate economic and philanthropic goals [12,17]. Though the marketing literature defines corporate hypocrisy as "the belief that a firm claims to be something that it is not" [1] (p. 79), the term hypocrisy is generally used to refer to "a pretense of having a virtuous character, moral or religious beliefs, principles, etc., that one does not really possess" [18]. Thus, in working language, hypocrisy refers to a perception regarding inconsistencies in outward behaviors or practices of an individual or institution. In this light, this study defines perceived corporate hypocrisy (PCH) as a perception of corporate pretensions of having a virtuous character. This perception could be created by all stakeholders, including consumers, investors, and employees. However, most research has focused only on consumers' and/or investors' perceptions, partly because such research was interested in marketing and brand management strategies [1,19]. This leaves us with a limited understanding of how employees react to PCH [3,20].

Employees, being important stakeholders [21], can also perceive corporate hypocrisy when their employers pretend to having virtuous characteristics [20]. To this end, it is critical to consider employees' negative perceptions and behavioral issues, as such concerns existing at the organizational and individual levels can impact not just on the immediate members of interest, but also on the entire supply chain network. Such perceptions may generate negative feelings towards the companies and their allies and affect employees' overall employment intentions [20]. Compared to consumers, employees have an insiders' view of their employer corporations' policies and practices. Therefore, their experiences of perceived corporate pretensions generating behavioral concerns, such as PCH, might be different than those experienced by consumers. Employees could be more exposed to inconsistent corporate policies and its execution, adding to their vulnerability of PCH, much of which has not been investigated in previous literature.

The U.S. retail industry, in particular, is one of the most significant employment sectors in the country. While it offers large employment opportunities [21,22], the sector also has a bad reputation due to unethical labor practices, immoral marketing approaches, and a questionable commitment to sustainability [20]. This sector has also been explored in the past and reported for existing PCH among

employees for their managers and corporations [20], but little to no literature exists on the empirical evaluation of such perceptions.

A possible explanation for this limited research on employees' PCH is the lack of a relevant instrument for estimating such perceptions exclusive to employees in their work settings. The scale currently available to measure PCH measures that of consumers [1], whose experience may be different from employees; in turn, this scale was used to link CSR to consumers' PCH. Moreover, Wagner et al.'s scale only partially focuses on the conceptual definition of corporate hypocrisy (e.g., acting inconsistently with regard to their assertions, pretending to be something different), and does not include several critical salient components of the hypocrisy construct, such as biased evaluations of moral transgressions within corporations, or a lack of morality [23,24]. Finally, the scale has tautological and self-referencing items, which artificially increase internal consistency and generate strong reliability, which is recommended to be avoided for effective scale development [25].

Despite the fact that the literature related to hypocrisy in the psychological, organizational, and philosophical fields has developed several scales to measure employees' perceptions surrounding organizational integrity and/or morality [1,23–25], none of them offer a comprehensive measurement tool to successfully measure employees' PCH. Finally, the literature mostly studied the presence or absence of hypocrisy using experimental manipulations and focused on the assessment of the concept as a self-identified characteristic [23]. Given that people are more likely to judge others than identify their own characteristic [26], there seems to be a gap in how to measure PCH objectively and as a perception formed by others. Considering that PCH is a complex, multi-faceted concept [27], to our knowledge, no single scale of analogous concepts of hypocrisy seems to exist to measure employees' PCH. Thus, the aim of this study was to develop a comprehensive and effective scale to measure employees' PCH to facilitate further future research surrounding PCH and CSR. Considering the contradictory nature of the U.S. retail industry sector and prior literature exploring PCH among its employees, this study used U.S. retail employees for PCH scale development.

To meet the purposes of the study, a three-stage approach was conducted, namely (a) item generation and initial purification, (b) content validity judgment, and (c) psychometric evaluation, scale refinement and validation [28–30]. In the first stage, a deductive approach was taken to generate items related to PCH, using scales and items from the literature. In the second stage, this set of items was reviewed by experts to ensure that they represented PCH well. Finally, in the third stage, the refined item bank was administered to the target population to evaluate items' psychometric properties, validity, and reliability. For the first two stages of item generation and initial purification as well as content validity judgment, qualitative approaches were taken via a thorough binning and winnowing, unstructured interviews, psychometric expert revisions, and cognitive interviews. For the final stage, a self-reported survey was administered through an online platform to field-test the item bank. With iterative analysis of the item bank, the final PCH scale was developed, which was then tested for association with employees' intention to leave the corporation and their attitude towards the corporation to confirm nomological validity.

For a corporation to be truly sustainable, it needs to acknowledge the people-centric characteristic of the entire supply chain and resolve psychological and behavioral issues. There can be various psychological, behavioral, and physiological concerns related to employees at the organizational and individual levels, which can significantly affect corporations' sustainability achievements. This scale is the first to assess and measure employees' PCH of their employers. Previous research has established that employees' negative perceptions may impact their consequent behavior [31]; this can be particularly damaging for corporate reputations because employees are a credible source of corporate identity for other/external stakeholders [32]. It is important to accurately measure PCH, because the disregard of negative perceptions may lead corporate employers to design inaccurate management strategies (which may have contributed to these perceptions at the first place). The scale developed herein will help corporations to obtain reliable, valid, and comprehensive measures of

PCH, while also facilitating further research into the causes and effects of negative perceptions in the workplace, potentially encouraging a more ethical and sustainable business climate.

2. Literature Review

2.1. Key Salient Components of Perceived Corporate Hypocrisy

A review of the literature helped identify four key salient components of the PCH construct [1,33]. First, hypocrisy is described as a pretense of morality, with the intention to appear moral to others and to gain benefits [23]. Such a commitment to morality is important in reference to integrity [34]. Researchers noted that hypocrisy is one's use of pretentious morality to conform to perceivers' standards of a morality system [23]. Also, researchers described the importance of different morality values, and one's commitment to them, for explaining hypocrisy motivation [35]. In this light, other scholars described that while disparity or conflict is a core of hypocrisy, such disparity must indicate the expression of, or claim to, some moral values to generate hypocrisy [36]. Moral practices and commitment to morality have been well studied in reference to integrity, a concept opposite to hypocrisy [34].

Second, a psychological contract breach, an important concept related to integrity within the organizational literature [34], is also thought to describe PCH. Psychological contract breach has been well studied in the past as an inconsistency between employers' promises and actions that undermines employees' perceptions of their employers' integrity [37]. Studies agreed that leader hypocrisy is an antonym of managers' behavioral integrity and identified psychological contract breach as a key component of leader hypocrisy [37]. However, psychological contract breach's contribution to hypocrisy perceptions has not been identified in the past.

Third, some studies consider PCH to form when one holds different standards of judgments for him/herself than for others [24]. It refers to when a discrepancy exists between how one believes others should behave in a given situation and how they themselves actually behave in the same or a similar situation [24]. Different from the previous concepts, which focus on actors' self-hypocritical behaviors, the double standard dimension points out the difference between actors' judgments of others' behaviors and themselves. Studies suggested that people might maintain a double standard to rationalize their selfish behavior, and thus might be perceived as hypocrites [33,38].

Fourth, other works posit PCH as the hypocrisy created when one behaves in a contradictory manner to his or her beliefs while pretending to support them, i.e., not practicing what one preaches [23,39]. These studies discussed hypocrisy as the divergence between corporations' words and actions, values, and behaviors, such that corporations preach about caring for their employees, customer service, certain other values, etc. but do not implement those in their actions [28,34,40]. Therefore, the new employees' PCH scale should encompass each of these psychological aspects.

2.2. Two Levels of Perceived Corporate Hypocrisy

PCH could arise in two major levels in workplace environments: (a) the system or organizational level and (b) the individual actor or member level [20]. Reviewing the four key salient components, PCH at the system or organizational level could be explained by the Theory of Organizational Culture (TOC), and at the individual actor or member level by theories of action (TOA). The next section describes the key psychological components and the development of a conceptual framework for employees' PCH.

2.2.1. PCH at the Organizational Level: Theory of Organizational Culture

The theory of organizational culture (TOC) provides a suitable framework for identifying salient aspects of PCH. The TOC rests on the premise that organizations have their own values, beliefs, myths, and rituals as their cultural properties [41]. It suggests that the culture of an organization is made up of three interrelated components, namely a sociocultural system, cultural system, and individual actors.

The sociocultural system is composed of organizations' formal structures, strategies, management processes and policies, and components to support organizations' reality and functioning [41]. Therefore, sociocultural systems do affect organizations' formal structures and such a formal structure becomes part of the organizational culture. The cultural system consists of shared meanings and symbols, such as myths, ideologies, values, and multiple cultural artifacts, which demonstrate the expressive and affective aspects of an organization [41]. Finally, the theory describes that the individual actors, with their experiences, personalities, characteristic features, leadership roles, and status, also contribute to the organizational culture [41].

According to the TOC, when an organization's structures, policies, and formal strategies align with its shared and accepted beliefs over time, with the individual actors interacting between these two systems, a strong culture develops within an organization [41]. Conversely, when these three components stand at odds with each other, the organizational culture becomes weakened. Most importantly, PCH may emerge from such dissonance to negatively impact employee performance and their commitment towards their organizations. Therefore, knowing how PCH could be potentially created within such organizational culture is especially important.

With this understanding of TOC, two of the four key salient components of hypocrisy were compared and reviewed against TOC. First, the perceived lack of morality that employees might hold against their employers, one of the key psychological components of PCH, seemed to be related to the overall cultural system of the employer organization. For example, the cultural system in TOC refers to shared meanings and symbols that demonstrate the expressive and affective aspects of an organization. A perceived inconsistency between these messages, shared meanings conveyed among people and employees (to achieve cohesion and collaboration), or a perceived inconsistency between those values used for structuring tasks and goal achievement can lead employees to question the corporations' morality. Such inconsistency between corporate ideologies and values, and its dominant actors, can create perceptions of a lack of morality among the employees and create room for PCH [41]. In this light, the perceived lack of morality was deemed to be a key salient component of PCH.

Second, the psychological contract breach corresponds with the sociocultural system of TOC. Given that the sociocultural system refers to organizations' formal strategies, reward and motivation system, psychological contract breach was thought to be influenced by the needs and preferences of their members [41]. Perceived inconsistency between employers' promises and actions would potentially generate PCH amongst employees. Thus, employees' psychological contract breach was hypothesized to be salient to PCH.

2.2.2. PCH at the Individual Level: Theories of Action

While the TOC offers a grand framework to explain complex organizational life and behaviors, it does not explain why such organizational behaviors take place. The theories of action (TOA; [42]) gives us insights into these questions. TOA illuminates the discrepancy between one's belief and behavior, which could potentially be associated with PCH [42], and explains the psychology behind organizational behaviors using two components: espoused theories and theories-in-use [42].

Espoused theories are the theories that one claims to follow explicitly, often reflected by values or views made to appeal to the target audience. Theories-in-use are the governing theories that support one's actions and eventually what is manifested in one's behaviors. While a congruence between these two components of TOA often influences the effectiveness of action, incongruence can be related to PCH [42]. The gap between what organizations say and what they do, which leads to the development of PCH among observers, is reflected in the disconnection between organizational espoused beliefs, values, principles, and the allegiance given to these beliefs in communication and the actual organizational behavior [34,42]. Therefore, understanding how organizations design their behaviors, how such plans and behaviors can be explained by the TOA, and how PCH could potentially be created due to the gaps between espoused and used theories are important questions to explore while developing a scale for PCH.

With such understanding of TOA, the remaining two of the four key salient components of hypocrisy were compared and reviewed against TOA. Third, the double standard nature of hypocrisy seems to occur at the individual level, which is related to the individual actors of corporate culture. Given that actors, with their experiences and status, contribute to the organizational culture [41], according to TOC, one's double standard in favoring him/herself over others can be influenced by the actors' personalities or organizational roles. This viewpoint was consistent with another study that argued that hypocrisy is perceived when people tend to hold different levels of moral standards for themselves compared to others, such that they tend to be more lenient towards themselves [24]. So, when corporations use different criteria to judge different groups of employees, or when they are more favorable towards certain employee(s) over others, or when they set different (and more forgiving) standards for themselves than for their employees, such double standards might create room for PCH. Previous PCH research also found employees reporting their corporations and supervisors for double standards, i.e., corporations using different standards of judgments for different people [20]. Thus, employees' perceptions that their employers might have different standards for different people, i.e., perceived double standards, would be salient to PCH.

Fourth, the employees' perceived gap between employers' value and behavior, or the value-behavior gap, was explained by TOA. For example, espoused theories, reflected by values a corporation wants to communicate, may contradict the theories-in-use reflected by corporate members' actions and actual behaviors [42], perhaps creating a value-behavior gap. That is, an inconsistency can exist between managers' values/behaviors and organizational values, thus leading to a value-behavior gap. The literature suggests that when people are seen to behave in a way that is contradictory to their statements and beliefs, but pretend to support such beliefs, that can be perceived as hypocrisy [16]. Even within an organizational setting, such divergence between espoused values, as expressed by words and enacted actions, is unanimously judged by others as hypocrisy [34]. For example, a corporation itself may behave in a manner contradictory to its announced values and such inconsistencies might create a value-behavior gap. Or, a corporation may indicate actions that are consistent with their values, but these might not be enacted by the management, leading to perceptions of a value-behavior gap among employees. Considering that managers and supervisors are often viewed as the faces of corporations and, as such, they play important roles in employees' perceptions [43,44], managers' inconsistent actions can lead to a value-behavior gap being noticed by employees, specifically in the retail industry [20]. Thus, it was proposed that employees' perceived gap between their managers' espoused values and behaviors, the value-behavior gap, would be salient to PCH among employees (see Figure 1).

It is important to note that such value-behavior inconsistencies are not always an issue of concern and, in fact, at certain times, such inconsistencies might have several benefits. Existing literature on PCH debates whether such inconsistencies are always to be avoided or whether inconsistencies can at times be beneficial for addressing stakeholders' contradictory needs [45]. However, such discussions are beyond the scope of the current study. Hence, this study focuses on the negative effects of such inconsistencies [16], i.e., as an aspect of PCH, with the notion that PCH can be harmful for the corporation [20].
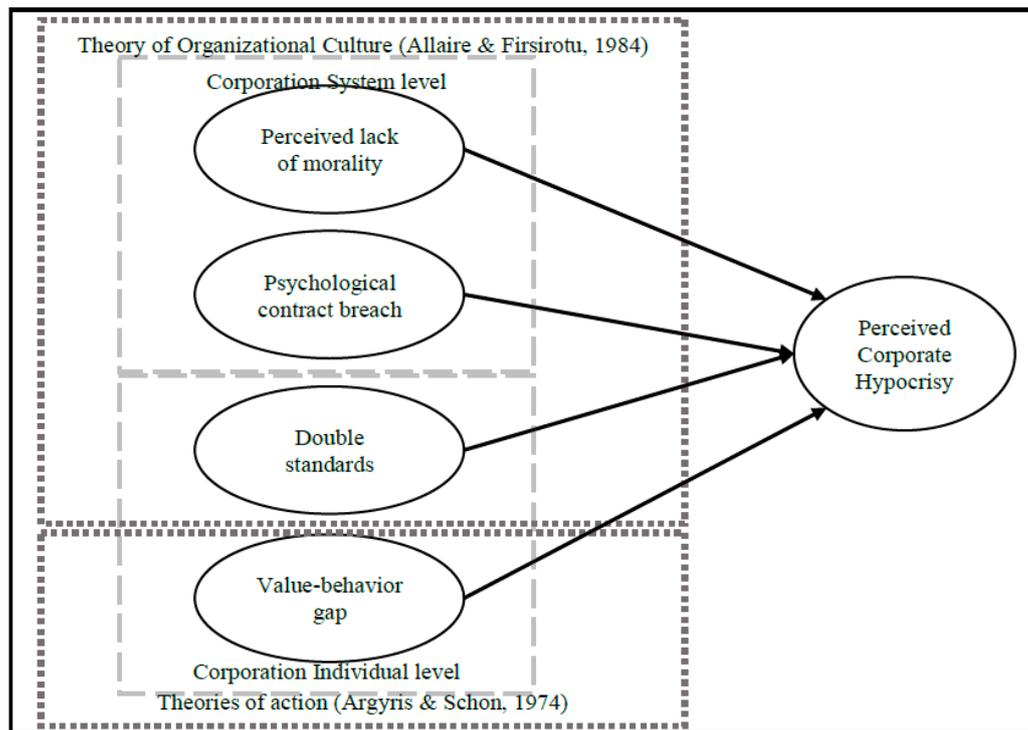
Figure 1. The conceptual model indicating perceived lack of morality, employees' psychological contract breach, perceived double standards, and value-behavior gap, as applied to PCH among employees.

3. Materials and Methods

Utilizing the conceptual model as the framework of this research, the well-established scale development procedures, as recommended in the literature, were followed [29–31]. A three-stage approach was conducted, namely (a) item generation and initial purification, (b) content validity judgment, and (c) psychometric evaluation, scale refinement, and validation. In the first stage, in study 1, a deductive approach was taken to generate items related to PCH [33]. The relevant scale items measuring the four proposed salient aspects of PCH were collected, adopted, and adapted from the literature to generate an initial item pool in this study. In the second stage, in study 2, the PCH items were reviewed and assessed by employees from different corporations and scholarly experts, to ensure that all items represented the structure of PCH well [31]. Finally, in the third stage, study 3 utilizes a broad employee population to assess and provide evidence of the items' psychometric properties, reliability, and to validate the measurement model [31]. See Table 1 for the stages in the scale development. All of these phases were conducted in 2016–17. All subjects gave their informed consent before they participated in the study. The study was conducted in three stages in accordance with the Declaration of Helsinki, and the protocol was approved by the IRB (Project numbers 2008118, 2008119, and 2008369).

Table 1. A summary of the three stages of scale development in this study.

Stage	Aim	Approach	Research Method	Sample	Results	Action Taken	Resulting Number of Items
One (Item Generation and Initial Purification)	Item Generation	Qualitative	Literature review	-	Generation of pool of dimensions and items	Specification of construct domain	145
	Binning and Winnowing	Qualitative	Concept mapping	-	Items that did not seem to match the designated domains, were inconsistent with PCH definitions, redundant in nature, confusing to understand, narrowly generalizable, and overly context-specific were identified	Problem items were deleted and strong potential items were identified	74
Two (Content Validity Judgment)	Content expert validation	Qualitative	Unstructured interviews	9 retail employees	Items too broad in nature, confusing and overlapping in content were identified	Problem items were deleted	49
	Item Revisions	Qualitative	Review by experts	2 business professors	Lack of clear instructions and a relevant frame of reference	The instruction was more clearly framed, and a time period of reference was included. the number of overall items were not further reduced	49
	Cognitive interviews	Qualitative	Verbal probing	8 retail employees	Lack of clear instructions: (a) regarding how participants should respond if they have retail experiences in more than one corporation, and (b) for participants who may not have experienced all of the 49 negative experiences as suggested by the items	New instructions were written, and the items were likewise modified	49
Three (Psychometric Evaluation, Scale Refinement and Validation)	Psychometric evaluation, scale refinement, and validation	Quantitative	Self-reported survey	520 retail employees	Item bank violated assumptions of unidimensionality and local independence, had items cross loading on multiple dimensions, had poor item and model fit, and several uninformative items	Six iterative item reduction for scale refinement	9

Throughout the scale development, researchers used the U.S. retail industry as the sample population framework. It is one of the significant employment sectors in the country, with an increase of 229,000 in employment from 2016 to 2017 [21,22]. At the same time, while the industry offers various entry-level jobs with little or no requirement of special skills and is often a popular choice for employment, the sector is also criticized as a 'bad' and potentially hypocritical industry [20,46]. For example, false promotion is common in the textile and apparel retailing sub-sector [47]. In 2012, New Balance (a U.S. footwear manufacturer and retail corporation), and, in 2016, Lord & Taylor (a U.S. luxury products retail corporation) were charged with misleading advertisements, and these companies were subjected to civil penalties and corrective actions per the Federal Trade Commission [48,49]. False and inaccurate claims and mislabeling are other common unethical issues in the retail sector [10], potentially creating PCH. Therefore, the retail sector was thought to offer room for detecting and measuring PCH as noticed by employees [20].

3.1. Stage 1: Item Generation and Initial Purification

The relevant scale items measuring the four proposed salient aspects of PCH were collected and adopted in this study to generate an initial item pool. The processes of binning and winnowing were related to the initial purification of the item pool in multiple steps [50]. Researchers used binning to identify and remove any redundant items from the pool and also to identify strong potential items conforming to the previous literature [50]. Winnowing was used to reduce the number of items within each bin, retaining only the items that were most relevant to the construct [50].

3.2. Stage 2: Content Validity Judgment

In this second stage of the study, three qualitative steps, namely (a) content expert validation, (b) item revisions and (c) cognitive interviews, helped to establish that the items matched expectations about the structure of the latent trait of interest [50]. In a scale development study, careful and systematic qualitative approaches are recommended to help assure the content validity of the generated item bank [51].

The purpose of content expert validation was to ensure content validity by capturing individuals' relevant PCH-related experiences [51]. A small group of retail employees ($N = 9$), representing varying retail industries, job ranks, and demographics, participated in this study to assess the reduced pool. The experts participated in unstructured interviews, answering questions about items, items' vocabulary, the perceptions those items might generate, and their experiences related to PCH. During this process, theme saturation occurred [52], indicating that the collected data are sufficient for analysis and that further new data might not generate any new findings.

The purpose of item revisions was to revise the item pool such that the measurement model reflects a consistent response format, certain literacy level expectations, fewer cognitive difficulties, and less ambiguous statements [50]. Considering that the item bank was generated from scales across various disciplines, such varieties could challenge the coherence of the item bank and generate questionable measures [50]. To achieve these objectives, two psychometric experts associated with a large university in the mid-Western USA were invited to participate in the study. The experts were business professors with scholarly experience in corporate behaviors and had no familiarity with this research. The experts reviewed the item pool, including the instructions for participants to answer those items, and gave specific feedback. At this stage, a four-point Likert scale response format with no neutral point was employed. Some researchers advocate using a forced-choice orientation as the item response format to force their participants to not take a middle ground on sensitive issues [29,53]. Thus, four response choices, from "strongly disagree" to "strongly agree," were designed so that a higher score indicated higher PCH amongst employees. While some scale development studies explain that negatively worded items may lead to artifactual or spurious responses from participants [54], other studies argue that no noticeably problematic pattern was found when using negatively worded items [55]. Additionally, applied to this study, which focuses on questionable morality, unfulfilled

promises, and the inconsistencies of a corporation, the use of negatively worded questions were deemed to be relevant and thus a requirement.

Cognitive interviews were conducted to ensure that respondents understand the items, understand instructions, and know how to reply to the items [50]. A verbal probing approach was taken, in which participating retail employees first answered the items as a regular survey. Later, they participated in follow-up structured interviews to discuss the items' comprehensibility, relevance of response categories, ease of understanding instructions, and overall layout [50]. Participants (*N* = 8) with more than a year of experience in varied retail industries responded to items using a four-point Likert scale from 1 (*strongly disagree*) to 4 (*strongly agree*) and participated in interviews wherein they systematically reviewed their basis for every response. Again, eight cognitive interviews yielded theme saturation [52].

3.3. Stage 3: Psychometric Evaluation, Scale Refinement, and Validation

In this third and final stage, a quantitative study was designed to test for items' psychometric parameters, assess items and scale reliability, and create a parsimonious scale [30]. It was also conducted to validate the derived scale [50]. A self-reported survey, including items derived in previous stages, was designed with two additional items that were used as attention checks to identify participants not paying attention to the instructions. A new sample of 520 adult respondents, representative of the 12 industries of the U.S. retail sector and with at least one year of experience in retail, participated in this study. Qualtrics, a market research firm, facilitated the data collection for this study. Data were collected over a week in April 2017. Table 1 shows the demographic information in detail.

Participants were asked to recall their experiences as an employee in their retail corporation and rate the corporation on the 49 items. Along with the PCH items, participants were also asked to complete a scale on employees' turnover intentions [34] for checking convergent validity and a scale on attitude towards the corporation [1] for checking discriminant validity.

Data Analysis

Data were analyzed using techniques from item response theory (IRT). IRT is a statistical framework for psychological measurement that models the probability of a given response based on a set of relevant items within a scale [56]. It provides the advantage of making scale items and their parameters invariant of the population and offers independence between the scale and its items, thus making scale construction more meaningful for organizational, behavioral, and psychological research [57]. Literature on psychological assessments recommends the IRT approach over classical scaling methods, such as Classical Test Theory (CTT), since the IRT model-based approach recognizes that measurement precision might not be constant for all people [58]. CTT traditionally uses analyses such as exploratory and confirmatory factor analyses to confirm dimensionality and the Cronbach's coefficient alpha to evaluate scale reliability [59], and bases it solely on the item statistics, i.e., atypical item difficulty values and low item discrimination for identifying poor quality items [60]. Different than CTT, a primary goal of IRT is to estimate the locations of respondents along a latent trait continuum (e.g., their degree of PCH), while identifying the underlying constructs of the trait and determining the scale's adequacy to measure it. Unlike CTT, IRT describes the association between participants' levels on a latent trait and the probability of a specific response to an item using a non-linear function [60]. It measures item parameters using logistic models instead of proportions and item-scale correlations; the parameters of each item are not assumed to have equal contribution and, therefore, are factored into the obtained score of each participant [61]. IRT analyzes the assumption of unidimensionality similar to the CTT, and uses a factor-analytic model to assess dimensionality and the model-data fit. To confirm that there are no noteworthy residual correlations, IRT commonly uses Q3 statistics [56,62]. Along with the fit at the scale level, IRT also evaluates the fit at an individual item level using the S-X2 statistic [63]. In IRT, the item characteristic curve (ICC) is used to analyze the probability and

identify items with good difficulty and discrimination parameters [59]. Also, category response curves (threshold parameters '*b*') are used to evaluate response options for each item and the likeliness of each response category being selected based on the participant's level of the latent trait. Different from CTT, in IRT, reliability is evaluated using item information (to assess every item's measurement precision for distinguishing between participants based on their levels of latent trait) and scale information (to assess the range of the latent trait the scale is most precisely informative about). Finally, for scale development using IRT, poor items are detected not only based on their discrimination and difficulty indices, but also using their goodness-of-fit, i.e., the analysis of their residuals. Items are selected based not only on their relevance to the scale, but also on the amount of information contributed by each item individually [60]. Based on the above, the process of data analysis proceeded with evaluating three assumptions, selecting and fitting a model to the data, and determining the fit at the scale and individual item levels [56].

IRT assumptions. We first tested three critical assumptions of IRT: (a) unidimensionality, (b) local independence, and (c) monotonicity [56]. According to the first assumption, all scale items should represent a single underlying latent trait, such as PCH, and item covariance should be attributable to the presence of PCH alone. Principal component analysis (PCA), inspecting dimensions' eigenvalues, ratio of the eigenvalue of the dominant dimension to the eigenvalues of the second and other subsequent dimensions, and the bend or 'knee' in scree plots, was conducted to check for and confirm unidimensionality [50,64]. Using the criterion of eigenvalue greater than 1 [65] threshold of a minimum of 20% variance explained by the first dimension, and a scree test to determine the optimal number of dimensions to be extracted, the assumption of unidimensionality was evaluated [66]. According to the second assumption, there should not be any residual correlations among the items after controlling for the effect of the dominant dimension. In this study, the Q_3 statistics [57] of all items of the item bank, represented by the lower triangle of the residual correlation matrix, were used to detect local independence violations. The recommended critical value of 0.2 was used to identify local dependence, such that any residual correlations among items above 0.2, after controlling for the dominant dimension, would indicate violation of the second assumption. The third assumption holds that the probability of endorsing an item should increase as the underlying level of the latent trait increases. The assumption of monotonicity was checked in this study by reviewing non-parametric IRT plots [67,68]. The relationships between PCH and its item responses, along with the probability of individuals answering those items based on their levels of PCH, need to be similar to a S-shaped curve [69].

Item parameters and model fit. Once the data were confirmed to meet these assumptions, an appropriate IRT model was fit to the collected data. The graded response model (GRM: [70]), which describes the probability of a participant endorsing a specific response category, was fit to the response data for item parameter calibration. To analyze the psychometric properties of the items, item characteristic curves (ICC) and various parameters (e.g., threshold, discrimination, and person location parameters) of the GRM were estimated and reviewed. ICCs for individual items were plotted where the *x*-axis represents a standardized PCH continuum ranging from −3 to +3, with 0 representing average PCH, and the *y*-axis represents participants' probability of selecting a given response option for the polytomous items of the test [56]. Thus, the ICCs displayed the probability of endorsing each particular response among the four response categories, conditional on the PCH of the respondent.

For polytomous (e.g., Likert-type) items, the threshold parameter '*b*' refers to the points along the latent trait continuum at which a higher response category becomes more likely than a lower one. In the present study, '*b*' parameters were analyzed to understand how participants with different levels of PCH selected among the given response categories for each item in the bank. The discrimination parameter '*a*' refers to an item's ability to differentiate among individuals with different levels of the latent trait [56]. A value of 0.01 to 0.34 is considered to represent very low discrimination, 0.35 to 0.64 to represent low discrimination, 0.65 to 1.34 for moderate discrimination, 1.35 to 1.69 for high discrimination, and 1.7 or above for very high discrimination [71]. The person location parameter (θ)

represents the degree to which a respondent exhibits the latent trait [56]. In this study, estimates of θ indicated individual differences in PCH and was reviewed from the ICCs.

In addition to evaluating the psychometric properties of each item, we also ensured that the chosen model fits well with the data. The fit of the GRM to the observed proportions of response was evaluated using the polytomous S-X^2 statistic [63], and items were reviewed or deleted if they exhibited significant item misfit ($p < 0.05$). Global model fit was assessed using the M_2 statistic as well as various categorical data versions of common factor-analytic fit indices (e.g., RMSEA, CFI, TLI, SRMR) [56,72]. Specifically, we hoped to achieve the recommended benchmarks for these fit statistics: RMSEA < 0.08, CFI > 0.95, TLI > 0.95, and SRMR < 0.05 [73].

Reliability. A central concept in scale development is to ensure reliability, that is, the degree to which the items and scales are free of measurement error [50]. In IRT, reliability is described in terms of the amount of information provided by each item about the respondent's true location along the θ scale. Such information is indicated by the Item Information Functions (IIFs) of individual items, which can be summed to find the Test Information Function (TIF) of the full scale. High reliability is then represented by highly peaked IIFs and a TIF that covers a meaningful range of the latent trait [56]. In this study, both the IIFs and TIF were reviewed to assess scale reliability.

Scale refinement. To create a parsimonious and interpretable instrument, it is important to identify the best and poorest performing items and take relevant actions [74]. Poor psychometric properties include high local dependencies among items pairs, cross-loading on multiple dimensions, item parameters above acceptable range, weak item and model fit, and misleading or otherwise inappropriate item content [50]. In this study, six iterative reductions were conducted to identify and eliminate the problem items.

To address these issues, items exhibiting local dependence violations ($Q_3 > 0.2$), low factor loadings (<0.3), or cross-loading on multiple dimensions criteria were flagged for further review [50]. Then, the content of each flagged item was reviewed to check for contextual overlap with other items, and redundant items were eventually deleted. The new set of reduced items was again screened for poor psychometric properties; this cycle was repeated five additional times until all items satisfied the relevant psychometric criteria. PCA was conducted to assess the assumption of unidimensionality of the item pool, and such analysis was repeated for all five iterations. It is important to note that, although certain items were identified as exhibiting poor statistical properties, such items were not deleted if they were deemed to have important contextual contributions, thereby maintaining content validity.

Validation. To validate the PCH scale, content, convergent, discriminant, and nomological validity were assessed. Content validity (i.e., the adequacy with which the content of the final scale covers and represents the theoretical constructs of PCH) was established during the qualitative steps of item generation, as discussed previously. Internal structure validity of the final measure was established via evidence from PCA and GRM analyses. The goodness-of-fit criteria and the significance and direction of factor loadings were used to establish convergent validity of the PCH scale [29]. Discriminant validity was established by comparing the Average Variance Extracted (AVE) with the squared correlation between the constructs [75]. AVEs were calculated for each of the four theoretical domains that constituted PCH and such square root of such variances were compared with the Pearson bivariate correlations between those domains. Discriminant validity was achieved when the square root of AVEs were greater than the correlations [75].

Nomological validity was confirmed by examining whether the PCH scale correlated with theoretically related yet distinct concepts [76]. To this end, literature suggests attitude towards corporation [19] and turnover intention [34] as valuable constructs for assessing nomological validity. When forming an attitude towards their corporations, employees use their perceptions about corporations [19] both at the individual and organizational level. Thus, it was hypothesized that employees' PCH for their employers would be negatively related to their attitude towards corporation. Additionally, taking an approach similar to previous research [19], the nine items of the PCH scale and the four items measuring attitude towards corporation were concatenated into a single

scale. Then, a confirmatory factor analysis (CFA) was conducted in which a two-factor model was imposed on the concatenated scale; if the constructs are distinct, then a two-factor model will fit well. Similarly, other research suggests that employees' PCH plays a role in their intention to continue their employment in those companies, such that, the more employees perceived hypocrisy, the higher their turnover intentions [34]. Thus, employees' PCH was hypothesized to be positively related to their turnover intentions.

Differential item functioning. Finally, in scale development, it is important that items are free of any demographic bias and thus fair to all respondents. To ensure test fairness, IRT provides a powerful tool called differential item functioning (DIF) to confirm that people with similar levels of PCH generate similar test results, irrespective of any age, gender, race, education, and/or cultural differences [50]. Any items that exhibit a high degree of DIF are recommended for further review, such as inspection of item design, or deletion if any reconcilable issues found [56]. In this study, DIF was evaluated for participants' gender, age, and workplace, using the likelihood ratio (LR) χ^2 test as the detection criterion (at $\alpha = 0.05$), and McFadden's pseudo-R^2 as the magnitude measure [56].

4. Results

4.1. Stage 1: Item Generation and Initial Purification

The result of item generation in stage 1 was a 145-item pool that could potentially measure employees' PCH. Within this pool, 44 items measured perceived lack of morality (e.g., "My corporation is unfair" [77]); 33 items measured employees' psychological contract breach (e.g., "So far, my corporation has done a very bad job of fulfilling its promises to me" [78]); 16 items measured double standards (e.g., "My company's values often CHANGE when it comes to getting things done" [34]); and 47 items measured the value-behavior gap (e.g., "How my manager represents himself/herself to the public is very different from what happens internally" [79]).

The binning process started with 145 items, which were grouped in "bins" according to similarity of content [50]. Four bins were created; items that did not seem to match the designated bins were set aside for further review and eventually deleted. With winnowing, the item pool was reviewed to remove items inconsistent with PCH definitions, redundant in nature, confusing to understand, narrowly generalizable, or overly context-specific [50]. Thus, the systematic process of binning and winnowing helped to purify and reduce the pool of 145 items to the 74 items that offered the most accurate representation of PCH.

4.2. Stage 2: Content Validity Judgment

The content expert validation participants assessed the reduced pool of 74 items. Table 2 shows demographic information of the participants. Participants were mostly aged between 21–30, had more women than men, represented mostly Caucasian and African-American ethnicity, were mostly part-time employees and most of them worked at the retail shop floor. Participants noted that the experiences represented by the items were relevant to their practical experiences and perceptions about their employer corporations, and no additional construct was deemed necessary to reflect PCH. Also, five items were identified as too broad in nature, six items were identified to be confusing, and 18 items were thought to be overlapping in content. Finally, participants indicated the need to highlight the negative connotation of the items, and recommended capitalizing the negative words for added emphasis. This evaluation phase deleted and rephrased several items, resulting in a reduced item pool of 49 items.

Table 2. Sample descriptions for two qualitative steps of stage 2 and quantitative psychometric evaluation step of stage 3 of the study.

Variable	Stage 2 (Content Expert Validation)		Stage 2 (Cognitive Interview)		Stage 3 (Psychometric Evaluation)	
	N	%	N	%	N	%
Total Sample Size	9	-	8	-	520	-
Age						
18–20	0	0.00%	0	0.00%	28	5.38%
21–30	6	66.67%	7	87.50%	184	35.38%
31–40	2	22.22%	0	0.00%	150	28.85%
41–50	0	0.00%	1	12.50%	82	15.77%
51–60	1	11.11%	0	0.00%	62	11.92%
61 and above	0	0.00%	0	0.00%	14	2.69%
Gender						
Man	4	44.44%	2	25.00%	137	26.35%
Woman	5	55.56%	6	75.00%	378	72.69%
Prefer not to disclose	0	0.00%	0	0.00%	5	0.96%
Ethnicity						
Caucasian	3	33.33%	6	75.00%	392	75.38%
Hispanic	1	11.11%	0	0.00%	47	9.04%
African-American	3	33.33%	2	25.00%	39	7.50%
Asian	0	0.00%	0	0.00%	15	2.88%
Other	1	11.11%	0	0.00%	27	5.19%
Employment Status						
Part-time	5	55.56%	6	75.00%	172	33.08%
Full time	4	44.44%	2	25.00%	331	63.65%
Not employed	0	0.00%	0	0.00%	9	1.73%
Retired	0	0.00%	0	0.00%	8	1.54%
Workplace						
Retail shop floor	7	77.78%	6	75.00%	417	80.19%
Corporate office (on-site and off-site)	2	22.22%	2	25.00%	103	19.81%

In item revisions, the two participants reviewed the item pool of 49 items and gave specific feedback regarding the lack of clear instructions for participants and the need for including a PCH relevant frame of reference in the instructions. Thus, based on the experts' comments, while the number of overall items were not further reduced, the instruction was more clearly framed, and a time period of reference was included for participants to refer to when providing responses to the items.

Analysis of the cognitive interviews indicated: (a) a lack of clear instructions regarding how participants should respond if they have retail experiences in more than one corporation, and (b) a lack of clear instruction for participants who may not have experienced all of the 49 negative experiences as suggested by the items. Participants represented demography similar to that of content expert validation participants. Participants were mostly aged between 21–30, had more women than men, represented mostly Caucasian ethnicity, were mostly part-time employees and most of them worked at the retail shop floor. Table 2 shows demographic information in detail. Based on the interview results, new instructions were written, and the items were likewise modified. Thus, a bank of 49 items was developed for testing in the next stage (see Table 3).

Table 3. The reduced item pool as derived after stages 1 and 2 of the study.

Number	Items
1	My company BREAKS most of the promises made during recruitment.
2	My company breaks many of its promises to me for NO fault of my own.
3	My company mostly FAILS to meet its obligations to me.
4	My company often does NOT fulfill its most important obligations to me.
5	My company often LIES to me.
6	My company does NOT acknowledge employees as humans.
7	My company often THROWS ME UNDER THE BUS for its own benefits.
8	My company has NO compassion for its employees.
9	My company uses my mistakes to INDIVIDUALIZE me.
10	My company TAKES CREDIT for my ideas.
11	My company RANDOMLY changes its goals without communicating this to employees.
12	My company does NOT have its employees' best interests at heart.
13	My company tends to look out only for ITSELF.
14	My company's policies do NOT match the promises made to employees.
15	My company makes promises to employees, which I can RARELY expect to actually happen.
16	My company is UNJUST to its employees.
17	My company is UNFAIR to its employees.
18	My company is SELFISH.
19	My company has almost NO moral principles.
20	My experience in my company is often NOT personally satisfying.
21	My company does NOT care for its employees, but only for money.
22	My company is INEFFICIENT in enacting its own set principles.
23	My company often COMPROMISES its important values as shared in public.
24	My company PRIORITIZES its benefits over employees' benefits.
25	My company PRETENDS to appear moral.
26	My company engages in morally WRONG acts when it can get away with them.
27	My company FIRES people on unjust grounds when it can get away with it.
28	My company's moral values are NOT the same as my moral values.
29	My company's values often CHANGE when it comes to getting things done.
30	My company does NOT behave honestly when dealing with employees.
31	My company does NOT behave ethically when dealing with employees.
32	My supervisor does NOT practice what (s)he preaches.
33	My supervisor GETS AWAY with doing things I can't.
34	There is an 'us' VERSUS 'him/her' between employees and supervisor.
35	My supervisor does NOT apply the same standards for performance to all employees.
36	My supervisor does NOT hold everyone at all levels equally accountable for their mistakes.
37	My supervisor does NOT give me enough authority to carry out my job responsibilities, but penalizes me for lack of performance.
38	The amount of work my supervisor requires me to do CONFLICTS with the quality of work (s)he expects.
39	My supervisor FAVORS employees based on her/his personal preferences rather than employees' abilities.
40	My supervisor does NOT conduct herself/himself according to the same values (s)he talks about.
41	My supervisor PRETENDS to be someone (s)he is not.
42	The values my supervisor communicates to the society are NOT consistent with employees' experiences at work.
43	The way my supervisor represents himself/herself to the public is very DIFFERENT from what happens internally.
44	There is a DIFFERENCE between what my supervisor says and what (s)he does.
45	My supervisor's behaviors do NOT reflect the company's values.
46	My supervisor MISLEADS employees with her/his communication and conflicting actions.
47	My supervisor is DECEPTIVE.
48	My supervisor shows employees what they want to see INSTEAD of the reality of the situation.
49	My supervisor MISLEADS employees about the real motives of the company.

4.3. Stage 3: Psychometric Evaluation, Scale Refinement, and Validation

4.3.1. Demographic Characteristics

A descriptive analysis of the data was conducted to understand the demographic characteristics of the study participants. Majority of the participants (35.5%) were aged between 21 and 30 years, 72.7% participants were women and 75.2% participants identified themselves as Caucasian. Majority of the participants (63.7%) were employed full time (40 or more hours/week) and 80.19% were representing employees working on the retail shop floor. Table 1 shows demographic information in detail.

4.3.2. Initial Item Bank

The full item bank yielded five principal dimensions, accounting for 68.97% of the total variance. Using the criterion of eigenvalue > 1 [65], the full item bank yielded five principal dimensions with eigenvalues 25.29, 4.21, 2.04, 1.14, and 1.10, and the data were observed to violate the assumption of unidimensionality. In the lower triangle of the Q_3 matrix, 143 item pairs exhibited high residual correlations beyond the critical value of 0.2, thus indicating local dependence in violation of the second assumption of IRT [80]. The non-parametric IRT plots increased monotonically for all 49 items, thus satisfying the assumption of monotonicity [81].

The ICCs based on the GRM indicated that all response categories across the complete bank were endorsed, signifying that all 49 items included an adequate number of response categories. Item discrimination '*a*' parameter of the full bank ranged from 0.89 to 3.09, indicating that the items were successful in differentiating among participants with varying levels of PCH. Given all the items had four response categories, three threshold '*b*' parameters were observed for each item, with an overall range from −3.38 to 0.978. All the items' higher threshold parameter estimates were close to or slightly higher than 0, indicating that participants with lower or close to average levels of PCH were prone to select the highest response options.

Consider, for example, item 29 ("*My company's values often CHANGE when it comes to getting things done*") from the initial item bank. This item discriminated well (*a* = 2.091) among participants and the threshold parameters (b_1 = −1.776, b_2 = −0.961, b_3 = 0.561) showed the points where the category response curves intersect. Participants with PCH of 0.56 standard deviations above average or higher were likely to select "Strongly agree" on Item 29. All the items' higher threshold parameter estimates were close to or slightly higher than 0, indicating that the item bank captured mostly lower and average levels of PCH (see Figure 2).
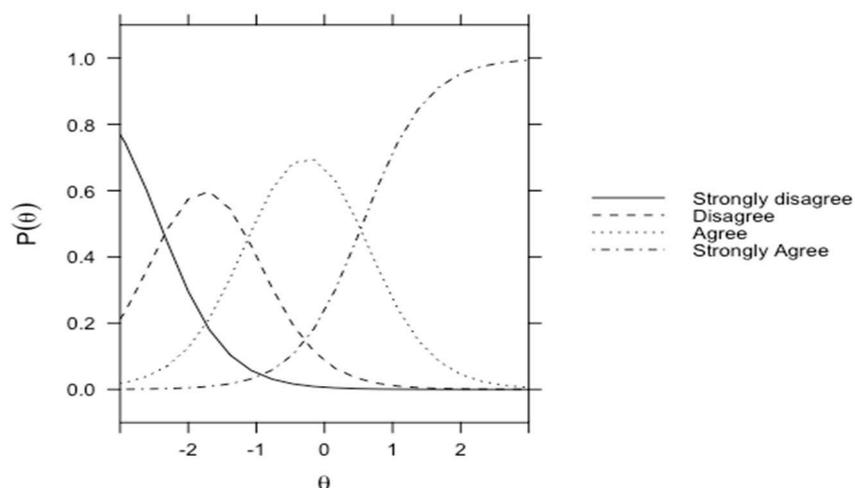

Figure 2. This figure shows an example Item Characteristic Curve (ICC) from the initial item bank. ICC of item 29 along with the probabilities for all response categories.

At the scale level, multiple fit indices revealed poor fit of the full bank to the data (M_2 (1029) = 8795.423 $p < 0.01$, RMSEA = 0.12, SRMR = 0.13, CFI = 0.91, and TLI = 0.90). However, the S-X^2 item fit statistic revealed that all but three items were fit well ($p > 0.05$) by the GRM.

The IIFs for certain items in the lack of morality, double standards and value-behavior gap domains had lower peaks, indicating that these items were less reliable. Figure 3, for example, presents the IIFs of items 7 and 17 from the full bank analysis. The IIFs of both items positioned over the lower end of the scale, meaning these items offer their most precise measurement when the respondent is low in PCH. However, the IIF for item 7 ("My company often THROWS ME UNDER THE BUS because of its own benefits") has a lower peak than that of item 17 ("My company is UNFAIR to its employees"), meaning item 7 offers less information than item 17 about the respondent's true level of PCH.
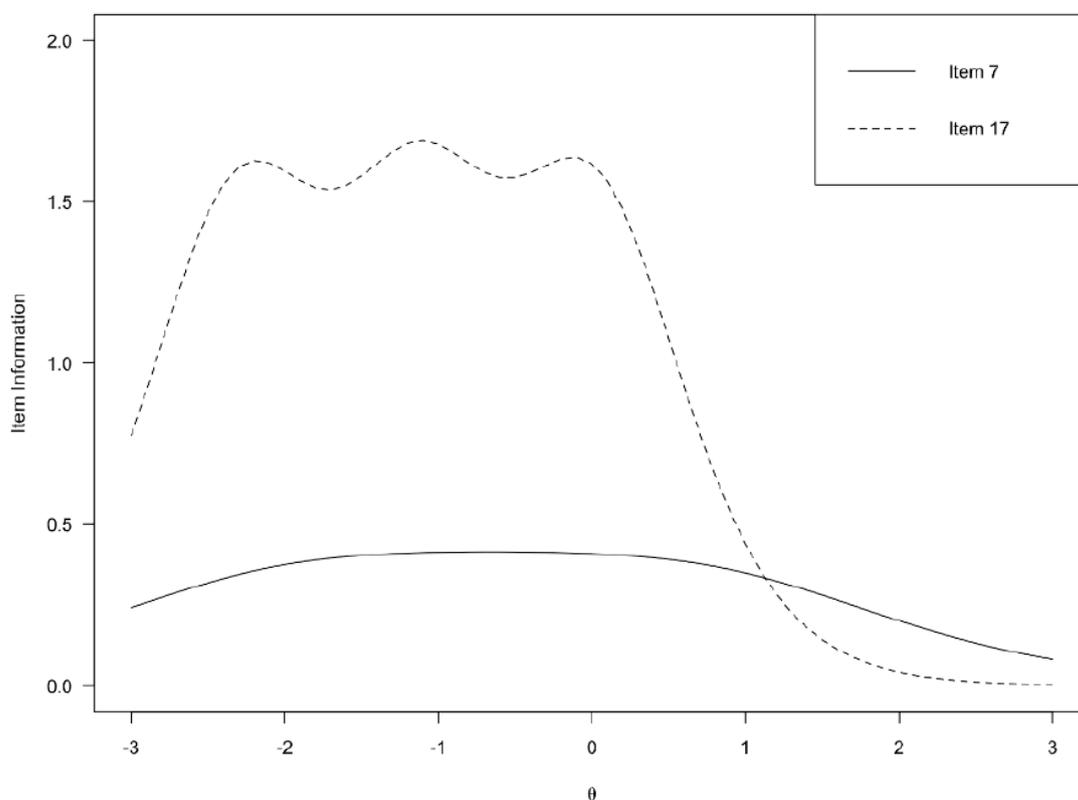

Figure 3. This figure shows an example comparison of Item Information Functions (IIFs) from the initial item bank. IIFs of items 7 and 17, indicating a lower and higher peak.

Based on the analyses above, it was determined that the initial 49-item bank violated the assumptions of unidimensionality and local independence, and had items cross loading on multiple dimensions, poor model fit, and several uninformative items. Thus, it was subjected to iterative item reduction to identify and remove all problem items, refine the scale, and develop a more parsimonious scale.

4.3.3. Scale Refinement

The PCH item bank was subjected to six iterative reductions using the recommended criteria. The results of these iterations are summarized in Table 4.

The most psychometrically problematic items from the full bank were removed and the reduced item set was subjected to each of the statistical criteria described above. This process was repeated until we identified a set of items that met all data assumptions, exhibited strong psychometric properties, and demonstrated high reliability. A total of six such iterations were completed, eventually uncovering a nine-item scale, which we refer to hereafter as the PCH scale.

Table 4. Summary of scale refinement performed over six iterative reductions in stage 3 of the study.

Iterations	Items Deleted	Number of Items	Number of Dimensions Extracted	M_2	RMSEA	SRMR	CFI	TLI
Initial item bank	-	49	5	M_2 (1029) = 8795.42, $p < 0.001$	0.12	0.13	0.91	0.90
1	1, 5, 9, 24, 45	44	4	M_2 (814) = 7426.25, $p < 0.001$	0.12	0.12	0.92	0.91
2	2, 3, 4, 6, 7, 18, 27, 33	36	3	M_2 (522) = 5440.66, $p < 0.001$	0.13	0.11	0.89	0.89
3	8, 11, 12, 13, 15, 20, 23, 41, 44, 47, 47, 48, 49	23	2	M_2 (184) = 1307.10, $p < 0.001$	0.10	0.06	0.92	0.91
4	19, 35, 39, 43	18	1	M_2 (99) = 677.17, $p < 0.001$	0.11	0.05	0.91	0.89
5	34, 37, 38, 40	14	1	M_2 (49) = 233.15, $p < 0.001$	0.08	0.04	0.93	0.92
6	16, 22, 28, 30, 31	9	1	M_2 (9) = 19.44, $p = 0.02$	0.04	0.04	0.99	0.99

The PCH scale satisfied all IRT assumptions: PCA yielded a one-factor solution with eigenvalue of 5.7 explaining 63.24% of total variance, thereby affirming unidimensionality, while Q3 statistics and non-parametric IRT plots supported that the data met the assumptions of local independence and monotonicity. Table 5 displays the results of the GRM fit analysis, which demonstrated a good overall fit to the PCH data (M_2 = 19.44 ($p = 0.02$), RMSEA = 0.04, SRMR = 0.046, CFI = 0.99, TLI = 0.99). The factor loadings (derived from the IRT parameters) were strong and supported a single dimension (i.e., PCH). Regarding item fit, only the first item exhibited significant item misfit, though the S-X^2 value was barely below the $p < 0.05$ threshold. The discrimination 'a' parameter ranged from 2.03 to 3.35 and the threshold 'b' parameters ranged from −2.03 to 0.47, thereby indicating acceptable item characteristics. The ICCs of each item indicated all response categories were endorsed by the participants. Finally, regarding the reliability of the PCH scale, maximum measurement precision was associated with individuals of moderately low to slightly above average levels of PCH (−1.75 to 0.5), as shown in Figure 4. From the classical test theory perspective, the PCH scale exhibited high internal consistency reliability (Cronbach's α = 0.93).

Table 5. PCH scale and its final psychometric properties as a result of the scale refinement.

Items	Factor Loading	Percentage Variances	IRT Parameters				Item Fit		
			a	b_1	b_2	b_3	s-x2	df	p
1. My company's policies do NOT match with the promises made to employees.	0.77	0.592	2.581	−1.733	−0.719	0.243	41.774	28	0.046 *
2. My company is UNFAIR to its employees.	0.83	0.683	3.339	−1.654	−0.794	0.052	27.676	26	0.375
3. My company does NOT care for its employees, but money.	0.83	0.694	3.251	−1.726	−0.767	0.181	19.858	24	0.705
4. My company PRETENDS to appear moral.	0.84	0.700	3.347	−1.759	−0.818	0.284	31.017	23	0.122
5. My company engages in morally WRONG acts when it can get away with them.	0.82	0.680	3.358	−1.395	−0.625	0.191	31.959	27	0.234
6. My company's values often CHANGE when it comes to getting things done.	0.80	0.643	3.859	−1.723	−0.765	0.473	37.833	26	0.063
7. My supervisor does NOT practice what (s)he preaches.	0.72	0.515	2.033	−2.035	−1.237	0.063	33.423	31	0.350
8. My supervisor does NOT hold everyone at all levels equally accountable for their mistakes.	0.78	0.610	2.571	−1.981	−1.123	0.038	31.604	24	0.137
9. The values my supervisor communicates to the society are NOT consistent with employees' experiences at work.	0.76	0.583	2.465	−1.944	−0.921	0.346	26.230	27	0.506

* $p < 0.05$.
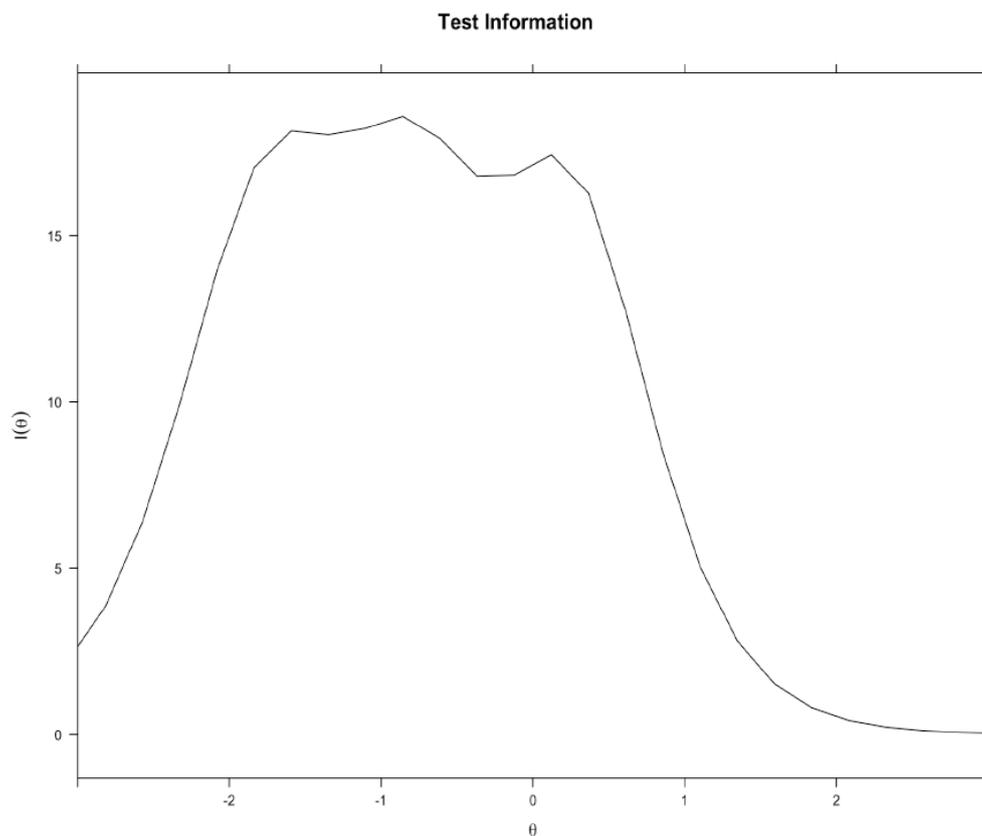
Test Information

Figure 4. Test information Function (TIF) plot for PCH scale, illustrating a higher curve at the left end of the scale.

4.3.4. Determination of Validity

The goodness-of-fit indicators for the PCH scale were adequate (CFI = 0.98, RMSEA = 0.04, M_2 = 19.44, *df* = 9, *p* = 0.02). Table 3 shows that the factor loadings for all nine items were positive, significant and above 0.7 in relation to PCH thereby indicating the convergent validity of the model. Similarly, all AVEs (ranging from 0.57 to 0.68) were above 0.5 providing good indications of convergent validity (refer to Table 5). As the AVEs were higher than the inter-construct correlations (between 0.25 and 0.51), discriminant validity was also implied (refer to Table 6).

Table 6. Discriminant validity assessment.

	Perceived Lack of Morality	Double Standards	Psychological Contract Breach	Value-Behavior Gap
Perceived lack of morality	0.82	0.51 **	0.5 **	0.44 **
Double standards		0.75	0.29 **	0.4 **
Psychological contract breach			0.77	0.26 **
Value-action gap				0.76

** *p* < 0.01. The diagonal represents the square root of AVEs for each of the four theoretical domains that constitute PCH. The table reports the correlation coefficient between these domains.

Regarding nomological validity, the PCH scale was significantly and negatively correlated to the attitude towards corporation scale, *r* = −0.45, *p* < 0.01, as hypothesized. The two-factor CFA fit the data well (CFI = 0.97, TLI = 0.96, RMSEA = 0.06, SRMR = 0.04), indicating that PCH is theoretically distinct from attitudes towards the corporation. Finally, nomological validity was established by a positive and significant correlation between PCH and employees' turnover intentions, *r* = 0.56, *p* < 0.01.

4.3.5. Differential Item Functioning

No DIF conditions regarding participants' gender were observed. However, four items were flagged for age-related DIF and one item was flagged for workplace-related DIF. All five DIF items were reviewed for content, wording, and response categories but none were found to be disputable. Additionally, all five items had small McFadden's pseudo-R^2 magnitude measures, meaning that the DIF was ignorable. Overall, the DIF analysis confirmed that the PCH scale is fair and unbiased with regard to gender, age, and workplace.

5. Discussion

Literature presents various cases to show that as corporations fail to act as they preach and to sustain an ethical culture, they often resort to acts to gain and improve their public perceptions and reputation [82]. Despite the existing literature on corporate hypocrisy, most studies investigated this concept from the perspectives of consumers [1,83], leaving a large gap in our understanding of employees' experiences related to corporate hypocrisy [3]. Such a gap is largely due to the lack of effective measurement tools for the construct, and therefore, the aim of this study was to develop scale for employees' PCH. By following the golden standards of scale development in IRT as well as methodology suggested by literature [30], this study developed a nine-item scale that is reliable and valid when measuring employees' PCH formed toward their corporations. The main contribution of this research is the development of the PCH scale with high reliability at Cronbach alpha of 0.93 and high level of content and construct validity.

5.1. Theoretical Contributions

The study findings make several important theoretical contributions to employee-related perceptions and hypocrisy research. First, this study filled a critical gap in the literature by constructing a scale specifically for employees' PCH by incorporating organizational, psychological, and philosophical literature related to hypocrisy. It brought comprehensive insights into employees' experiences and perceptions around corporate hypocrisy. So far, the literature has discussed several ways to measure corporate hypocrisy from the consumer's perspective [1]. This research sheds light on another important category of stakeholder: the employees. It uncovers how corporations can evoke such PCH in their employees and enables researchers to better measure employees' PCH towards their employers.

Second, four salient components were discovered and discussed in this study, which helps us understand employees' PCH: double standards, perceived lack of morality, breach of employees' psychological contract, and value-behavior gap. By doing so, this research identified PCH as a concept above and beyond behavioral inconsistencies, as mostly seen in organizational and marketing research [1,3] and achieved a more comprehensive and holistic measure of employees' PCH.

In this study, corporations were perceived as hypocritical when participants believed their corporations exercised different standards for the transgressions of different employees, and when they favored themselves over employees, i.e., had double standards. Therefore, this research supported the existing argument of double standards being a salient aspect of hypocrisy [26], adding to the evidence that such biased standards for transgressions exist and affect employees in their work setting. Another aspect of PCH was the perception that corporations were acting pretentiously when endorsing moral values or contradicting higher values with ulterior motives. This was a critical finding, considering that the current corporate hypocrisy literature [1,3,19] does not present any insight into the perceived lack of morality. Moreover, this research extended the PCH-related research, which has shown that perceptions of hypocrisy can be created by word-deed inconsistencies [1]. Based on these findings, this research proposes that future work on PCH needs to study the construct as a multifaceted one, explore the various salient aspects, and compare their contribution to hypocrisy perceptions and employees' psychological mechanisms.

Third, this research supports the theory of the organization culture and theories of action as the theoretical approaches to explain employees' PCH in different aspects. The TOC in literature has concentrated on how it affects job satisfaction or commitment and on employees' perceptions [84,85]. Similarly, the TOA had mostly been studied in the organizational literature regarding how dissonance can exist between espousals and enactments, and how managing such dissonance offer scopes of organizational learning [86,87]. Yet, little to no studies with TOA focused on how such dissonance might generate PCH among employees. This study is one of the first that looked at employees' PCH generated by potential dissonances in organizations. For future research, this strongly suggests using the TOC and TOA paradigms as they have good explanatory abilities for employees' PCH.

Finally, by using three systematic studies conceptualizing, developing and testing a comprehensive scale measuring employees' PCH, this scale was developed according to best practices recommended by IRT researchers, whereas only a minority of research follows such ideal standards [57]. The balanced application of qualitative and quantitative methods helped the study avoid over-reliance on statistical approaches while addressing every friction that arose in the process. Therefore, the PCH scale developed in this study is deemed to be robust and maintains methodological and social objectivity [88]. The study extends the usefulness of IRT in organizational research for a variety of applications. Using IRT to conceptualize and score employees' perceptions made the scale psychometrically sound, with the scale items and their parameters invariant of the population and independent, thus advancing the applicability of IRT.

5.2. Managerial Implications

The PCH scale is believed to provide managers with an effective and comprehensive measure of employees' PCH that addresses salient constructs that are absent from previous scales. A lack of understanding of employees' PCH and an inability to measure such perceptions can create counterproductive workplace behavior and challenge corporations in designing effective management strategies. Specifically related to corporate sustainability, when corporations' goals toward sustainability are manifested through other members' inconsistent behaviors, such lack of clear sustainable goals and corporate structures to achieve such goals can leave room for PCH among employees [11]. Considering the crucial linkage between CSR perceptions, sustainability and PCH, employees' negative perceptions can break corporate sustainability strategies, can lead to reduced work engagement, can trigger emotional exhaustion and lead to employee turnover [3]. With the PCH scale, any such perceptions can be timely detected and measured for corrective actions.

Second, this study found employees' PCH and their turnover intentions to be correlated. Thus, managers need to be able to understand and measure their employees' PCH in order to control it and reduce its potential influence on their turnover intentions. This might be particularly important in the retail sector, as it is challenged with higher turnover rates [46,89].

Third, our findings suggest that employees' PCH was related to their negative attitude towards their corporations. Negative perceptions toward corporations among their employees negatively affect the performance of corporate sustainability, i.e., corporations' sustainability performance decreases [11]. Also, employees can easily share such perceptions and related negative attitudes on various social media channels. Considering that employees are seen as credible sources and are believed to have an insider's view of the real corporate ethics and standards, their PCH and related negative attitudes might eventually harm corporations' overall sustainability performance and reputations. Thus, employers need to be able to assess their employees' PCH to address related potential problems. This scale would enable managers to do so.

Fourth, this study showed that PCH can be influenced by organizational culture. Such information calls for managerial attention to align business plans and strategies to shift corporate focus onto their ethical practices, values, and morality both at the system and at the individual level. Corporations must focus on having clear and aligned goals, and a well-defined corporate structure toward sustainability. This might be particularly important with the boom cycle of moral leadership, where employees expect

to find a deep sense of ethics, appreciation of core ideals, and pursuit of a higher purpose in their leaders [90].

6. Conclusions and Limitations

This research investigated PCH for employees and developed a statistically reliable and valid scale. The study identified four salient aspects of PCH by incorporating organizational, psychological, and philosophical literature related to hypocrisy. The review of literature indicated of these four aspects being crucial to explanation of hypocrisy perceptions among employees and were thus proposed to constitute the composite measure of PCH. Utilizing this conceptual framework and recommended scale development procedures, a deductive approach was taken as the fundamental step to generate an initial item pool. Established scales from literature that measured each of the four salient components individually, i.e., perceived lack of morality, perceived contract breach, double standards, and value-behavior gap, were taken as inspirational items and were either adopted directly or reframed in the PCH context.

While each such domain may seem to cover a lot more than corporate hypocrisy, the final scale empirically established every domain to be an important theoretically contributing element for employees' PCH. Thus, in line with the literature [28], the final scale of nine items, developed over six iterative scale refinement analyses, reported of PCH as a unidimensional complex construct and each of the items represented the four salient theoretical domains. The statistical dimensionality of the scale, along with other psychometric criteria such as item and model fits, were evaluated at each of the six iterations as criteria of scale refinement, and the iterations were stopped once the items reported sound statistical and conceptual information. Based on the above analyses, PCH was concluded as a unidimensional factor.

Comparing the results with that of the existing PCH scale for consumers [1], the scale for employees uniquely taps on experiences relevant to employees in their work settings, with items such as "My company's policies do NOT match with the promises made to employees" or "My supervisor does NOT hold everyone at all levels equally accountable for their mistakes." Different from Wagner et al.'s scale ($\alpha = 0.92$), this PCH scale reflects a more holistic conceptual definition of PCH but with comparable statistical reliability ($\alpha = 0.93$). Also, this scale reported on PCH both at the individual and the organizational level. Different to the scale of organizational hypocrisy [34], this PCH scale for employees empirically established that employees experience PCH at multiple levels within a corporation. The results support the theory of organizational culture as a structural foundation that explains employees' PCH among the sociocultural system, cultural system, and individual actors.

Limitations

Despite the valuable contributions this study could make, the study also has certain limitations related to external validity and test-retest reliability. First, the study focused on only retail corporations for scale development. Replicating the research applied to employees' experiences related to other corporate sectors, such as manufacturing and wholesale, are recommended to improve the generalizability of the results. Second, although several groups of participants were recruited at various stages of this study, all the study samples represented only U.S. employees, thus restricting the external validity of the study. Since employees' PCH related to their corporate employers might be of global interest, similar studies with participants representing other nationalities and cultures would be needed before the results can be generalized. This is because peoples' perceptions and reactions to different experiences might differ across countries and cultures. Finally, the study reported scale reliability based on item and test information. But another important form of reliability, the test-retest reliability, needs to obtained and such can be achieved by administering the scale multiple times over a period and analyzing the correlation between the scores can help establish the stability of the scale.

administration, S.G.; Resources, S.G.; Software, S.G. and W.B.; Supervision, S.G.; Validation, S.G. and W.B.; Visualization, S.G.; Writing—original draft, S.G. and J.H.-B.; Writing—review & editing, S.G., J.H.-B. and W.B.

References

1. Wagner, T.; Lutz, R.J.; Weitz, B.A. Corporate hypocrisy: Overcoming the threat of inconsistent corporate social responsibility perceptions. *J. Mark.* **2009**, *73*, 77–91. [CrossRef]
2. Aqueveque, C.; Encina, C. Corporate behavior, social cynicism, and their effect on individuals' perceptions of the company. *J. Bus. Ethics* **2010**, *91*, 311–324. [CrossRef]
3. Scheidler, S.; Edinger-Schons, L.M.; Spanjol, J.; Wieseke, J. Scrooge posing as Mother Teresa: How hypocritical social responsibility strategies hurt employees and firms. *J. Bus. Ethics* **2018**, 1–20. [CrossRef]
4. The Social Responsibility of Business is to Increase its Profits. Available online: http://umich.edu/~{}thecore/doc/Friedman.pdf (accessed on 15 March 2016).
5. First, Make Money: Also, Do Good. *The New York Times*. Available online: https://www.nytimes.com/2011/08/14/business/shared-value-gains-in-corporate-responsibility-efforts.html (accessed on 22 January 2016).
6. Is Corporate Social Responsibility Hypocritical Window Dressing? Forbes. Available online: http://www.forbes.com/sites/arjanschutte/2011/08/17/is-corporate-social-responsibilityhypocritical-window-dressing/ (accessed on 15 March 2016).
7. Cour, A.; Kromann, J. Euphemisms and hypocrisy in corporate philanthropy. *Bus. Ethics Eur. Rev.* **2011**, *20*, 267–279. [CrossRef]
8. Corporate Social Responsibilty: Is This Another Hoax? Available online: http://www.csringreece.gr/files/research/CSR-1289995213.pdf (accessed on 22 January 2016).
9. Diallo, M.F.; Lambey-Checchin, C. Consumers' perceptions of retail business ethics and loyalty to the retailer: The moderating role of social discount practices. *J. Bus. Ethics* **2017**, *141*, 435–449. [CrossRef]
10. Consumer Information. Available online: https://www.consumer.ftc.gov/blog (accessed on 22 January 2017).
11. Ha-Brookshire, J. Toward moral responsibility theories of corporate sustainability and sustainable supply chain. *J. Bus. Ethics* **2015**, *145*, 227–237. [CrossRef]
12. Kim, H.; Hur, W.M.; Yeo, J. Corporate brand trust as a mediator in the relationship between consumer perception of CSR, corporate hypocrisy, and corporate reputation. *Sustainability* **2015**, *7*, 3683–3694. [CrossRef]
13. Alcaniz, E.B.; Caceres, R.C.; Perez, R.C. Alliances between brands and social causes: The influence of company credibility on social responsibility image. *J. Bus. Ethics* **2010**, *96*, 169–186. [CrossRef]
14. Pivato, S.; Misani, N.; Tencati, A. The impact of corporate social responsibility on consumer trust: The case of organic food. *Bus. Ethics Eur. Rev.* **2008**, *17*, 3–12. [CrossRef]
15. Rupp, D.E.; Shao, R.; Skarlicki, D.P.; Paddock, E.L.; Kim, T.Y.; Nadisic, T. Corporate social responsibility and employee engagement: The moderating role of CSR-specific relative autonomy and individualism. *J. Organ. Behav.* **2018**, *39*, 559–579. [CrossRef]
16. Scheidler, S.; Schons, L.M.; Spanjol, J. Internal Marketing of Corporate Social Responsibility (CSR) Initiatives: CSR Portfolio Effects on Employee Perceptions of Corporate Hypocrisy, Attitudes, and Turnover. In *Rediscovering the Essentiality of Marketing*; Springer: Cham, Switzerland, 2016.
17. Fassin, Y.; Buelens, M. The hypocrisy-sincerity continuum in corporate communication and decision making: A model of corporate social responsibility and business ethics practices. *Manag. Decis.* **2011**, *49*, 586–600. [CrossRef]
18. Dictionary.com. Available online: http://www.dictionary.com/browse/hypocrisy (accessed on 22 January 2017).
19. Shim, K.; Yang, S.U. The effect of bad reputation: The occurrence of crisis, corporate social responsibility, and perceptions of hypocrisy and attitudes toward a company. *Public Relat. Rev.* **2016**, *42*, 68–78. [CrossRef]
20. Goswami, S.; Ha-Brookshire, J.E. Exploring US retail employees' experiences of corporate hypocrisy. *Organ. Manag. J.* **2016**, *13*, 168–178. [CrossRef]
21. National Retail Federation. Available online: https://nrf.com/resources/retail-library/the-economic-impact-of-the-us-retail-industry (accessed on 29 March 2017).

22. U.S. Bureau of Labor Statistics. Available online: https://www.bls.gov/news.release/archives/empsit_ 02032017.pdf (accessed on 29 March 2017).

23. Batson, C.D.; Thompson, E.R.; Chen, H. Moral hypocrisy: Addressing some alternatives. *J. Personal. Soc. Psychol.* **2002**, *83*, 330. [CrossRef]

24. Valdesolo, P.; DeSteno, D. Moral hypocrisy social groups and the flexibility of virtue. *Psychol. Sci.* **2007**, *18*, 689–690. [CrossRef] [PubMed]

25. Condon, D.M.; Revelle, W. The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence* **2014**, *43*, 52–64. [CrossRef]

26. Hale, W.J.; Pillow, D.R. Asymmetries in perceptions of self and others' hypocrisy: Rethinking the meaning and perception of the construct. *Eur. J. Soc. Psychol.* **2015**, *45*, 88–98. [CrossRef]

27. Alicke, M.; Gordon, E.; Rose, D. Hypocrisy: What counts? *Philos. Psychol.* **2013**, *26*, 673–701. [CrossRef]

28. Churchill, G.A., Jr. A paradigm for developing better measures of marketing constructs. *J. Mark. Res.* **1979**, *16*, 64–73. [CrossRef]

29. DeVellis, R. *Scale Development: Theory and Applications*; Sage Publications: Thousand Oaks, CA, USA, 1991.

30. Morgado, F.F.; Meireles, J.F.; Neves, C.M.; Amaral, A.C.; Ferreira, M.E. Scale Development. *Psicol. Reflexao e Critica* **2017**, *30*, 3. [CrossRef]

31. Flatt, S.J.; Kowalczyk, S.J. Creating competitive advantage through intangible assets: The direct and indirect effects of corporate culture and reputation. *J. Compet. Stud.* **2008**, *16*, 13.

32. Delmas, M. Stakeholders and competitive advantage: The case of ISO 14001. *Prod. Oper. Manag.* **2001**, *10*, 343–358. [CrossRef]

33. Philippe, T.W.; Koehler, J.W. A factor analytical study of perceived organizational hypocrisy. *SAM Adv. Manag. J.* **2005**, *70*, 13.

34. Simons, T. Behavioral integrity: The perceived alignment between managers' words and deeds as a research focus. *Organ. Sci.* **2002**, *13*, 18–35. [CrossRef]

35. Lönnqvist, J.E.; Irlenbusch, B.; Walkowitz, G. Moral hypocrisy: Impression management or self-deception? *J. Exp. Soc. Psychol.* **2014**, *55*, 53–62. [CrossRef]

36. Szabados, B.; Soifer, E. Hypocrisy, change of mind, and weakness of will: How to do moral philosophy with examples. *Metaphilosophy* **1999**, *30*, 60–78. [CrossRef]

37. Greenbaum, R.L.; Mawritz, M.B.; Piccolo, R.F. When leaders fail to "walk the talk". *J. Manag.* **2015**, *41*, 929–956. [CrossRef]

38. Monin, B.; Merritt, A. Moral hypocrisy, moral inconsistency, and the struggle for moral integrity. In *The Social Psychology of Morality: Exploring the Causes of 46 Good and Evil. Herzliya Series on Personality and Social Psychology*; Mikulincer, M., Shaver, P.R., Eds.; American Psychological Association: Washington, DC, USA, 2012; pp. 167–184.

39. Kouzes, J.M.; Posner, B.Z. *Credibility*; Jossey-Bass: San Francisco, CA, USA, 1993.

40. Abrahamson, E. Management fashion. *Acad. Manag. Rev.* **1996**, *21*, 254–285. [CrossRef]

41. Allaire, Y.; Firsirotu, M.E. Theories of organizational culture. *Organ. Stud.* **1984**, *5*, 193–226. [CrossRef]

42. Argyris, C.; Schon, D.A. *Theory in Practice: Increasing Professional Effectiveness*; Jossey-Bass: San Francisco, CA, USA, 1974.

43. Diamond, M.A. Dimensions of organizational culture and beyond. *Political Psychol.* **1991**, *12*, 509–522. [CrossRef]

44. Schein, E. *Organizational Culture & Leadership*; Jossey-Bass, John Wiley: San Francisco, CA, USA, 2010.

45. Brunsson, N. The necessary hypocrisy. *Thunderbird Int. Bus. Rev.* **1993**, *35*, 1–9. [CrossRef]

46. Williams, C.L.; Connell, C. "Looking good and sounding right": Aesthetic labor and social inequality in the retail industry. *Work Occup.* **2010**, *37*, 349–377. [CrossRef]

47. Time. Available online: http://time.com (accessed on 7 January 2016).

48. Federal Trade Commission. Available online: https://www.ftc.gov/news-events/press-releases/2016/03/ lord-taylor-settles-ftc-charges-it-deceived-consumers-through (accessed on 8 April 2016).

49. Huffington Post. Available online: http://www.huffingtonpost.com (accessed on 29 August 2012).

50. Revicki, D.A.; Chen, W.H.; Tucker, C. 16 developing item banks for patient-reported health outcomes. In *Handbook of Item Response Theory Modeling: Applications to Typical Performance Assessment*; Reise, S., Revicki, D., Eds.; Routledge: Abingdon, UK, 2014; pp. 334–363.

51. Brod, M.; Tesler, L.E.; Christensen, T.L. Qualitative research and content validity: Developing best practices based on science and experience. *Qual. Life Res.* **2009**, *18*, 1263. [CrossRef] [PubMed]

52. Spiggle, S. Analysis and interpretation of qualitative data in consumer research. *J. Consum. Res.* **1994**, *21*, 491–503. [CrossRef]

53. Craig, S.B.; Gustafson, S.B. Perceived leader integrity scale: An instrument for assessing employee perceptions of leader integrity. *Leadersh. Q.* **1998**, *9*, 127–145. [CrossRef]

54. Harvey, R.J.; Billings, R.S.; Nilan, K.J. Confirmatory factor analysis of the Job Diagnostic Survey: Good news and bad news. *J. Appl. Psychol.* **1985**, *70*, 461. [CrossRef]

55. Hinkin, T.R. A review of scale development practices in the study of organizations. *J. Manag.* **1995**, *21*, 967–988. [CrossRef]

56. De Ayala, R.J. *The Theory and Practice of Item Response Theory*; Guilford Publications: New York, NY, USA, 2009.

57. Foster, G.C.; Min, H.; Zickar, M.J. Review of item response theory practices in organizational research: Lessons learned and paths forward. *Organ. Res. Methods* **2017**, *20*, 465–486. [CrossRef]

58. Zickar, M.J. Modeling item-level data with item response theory. *Curr. Dir. Psychol. Sci.* **1998**, *7*, 104–109. [CrossRef]

59. Cappelleri, J.C.; Lundy, J.J.; Hays, R.D. Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clin. Ther.* **2014**, *36*, 648–662. [CrossRef]

60. Hambleton, R.K.; Jones, R.W. Comparison of classical test theory and item response theory and their applications to test development. *Instruct. Top. Educ. Meas. Ser.* **2012**, *12*, 38–47.

61. Hambleton, R.K.; Swaminathan, H.; Rogers, H.J. *Fundamentals of Item Response Theory*; Sage: Thousand Oaks, CA, USA, 1991.

62. Yen, W.M. Scaling performance assessments: Strategies for managing local item dependence. *J. Educ. Meas.* **1993**, *30*, 187–213. [CrossRef]

63. Orlando, M.; Thissen, D. Further investigation of the performance of S-X2: An item fit index for use with dichotomous item response theory models. *Appl. Psychol. Meas.* **2003**, *27*, 289–298. [CrossRef]

64. Funk, J.L.; Rogge, R.D. Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. *J. Fam. Psychol.* **2007**, *21*, 572. [CrossRef] [PubMed]

65. Kaiser, H.F. The application of electronic computers to factor analysis. *Educ. Psychol. Meas.* **1960**, *20*, 141–151. [CrossRef]

66. Cattell, R.B. The scree test for the number of factors. *Multivar. Behave. Res.* **1966**, *1*, 245–276. [CrossRef] [PubMed]

67. Mokken, R.J. *A Theory and Procedure of Scale Analysis*; De Gruyter: Berlin, Germany, 1971.

68. Sijtsma, K.; Molenaar, I.W. *Introduction to Nonparametric Item Response Theory*; Sage: Thousand Oaks, CA, USA, 2002.

69. Zanon, C.; Hutz, C.S.; Yoo, H.H.; Hambleton, R.K. An application of item response theory to psychological test development. *Psicologia Reflexão eCrítica* **2016**, *29*, 1–10. [CrossRef]

70. Samejima, F. Estimation of latent ability using a response pattern of graded scores. *Psychom. Monogr. Suppl.* **1969**, *34*, 100. [CrossRef]

71. The Basics of Item Response Theory. Available online: http://ericae.net/irt/baker (accessed on 15 November 2016).

72. Maydeu-Olivares, A.; Joe, H. Assessing approximate fit in categorical data analysis. *Multivar. Behav. Res.* **2014**, *49*, 305–328. [CrossRef]

73. Hu, L.T.; Bentler, P.M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* **1999**, *6*, 1–55. [CrossRef]

74. Gorsuch, R.L. Exploratory factor analysis: Its role in item analysis. *J. Personal. Assess.* **1997**, *68*, 532–560. [CrossRef] [PubMed]

75. Fornell, C.; Larcker, D. Evaluating structural equations models with unobservable variables and measurement error. *J. Mark. Res.* **1981**, *18*, 39–50. [CrossRef]

76. Malhotra, N.K.; Birks, D.F. *Marketing Research: An Applied Approach*; Prentice Hall: Upper Saddle River, NJ, USA, 2007.

77. Reidenbach, R.E.; Robin, D.P. Toward the development of a multidimensional scale for improving evaluations of business ethics. *J. Bus. Ethics* **1990**, *9*, 639–653. [CrossRef]

78. Robinson, S.L.; Morrison, E.W. The development of psychological contract breach and violation: A longitudinal study. *J. Organ. Behav.* **2000**, *21*, 525–546. [CrossRef]

79. Martinez, S.Y. Cultural integrity. Ph.D. Thesis, San Diego State University, San Diego, CA, USA, 2016.

80. Chen, W.H.; Thissen, D. Local dependence indexes for item pairs using item response theory. *J. Educ. Behav. Stat.* **1997**, *22*, 265–289. [CrossRef]

81. Van der Ark, L.A. Mokken scale analysis in R. *J. Stat. Softw.* **2007**, *20*, 1–19.

82. Barnett, M.L.; Salomon, R.M. Beyond dichotomy: The curvilinear relationship between social responsibility and financial performance. *Strateg. Manag. J.* **2006**, *27*, 1101–1122. [CrossRef]

83. Glavas, A. Corporate social responsibility and organizational psychology: An integrative review. *Front. Psychol.* **2016**, *7*, 144. [CrossRef]

84. Kangas, M.; Kaptein, M.; Huhtala, M.; Lämsä, A.M.; Pihlajasaari, P.; Feldt, T. Why do managers leave their organization? Investigating the role of ethical organizational culture in managerial turnover. *J. Bus. Ethics* **2016**, 1–17. [CrossRef]

85. Racelis, A.D. Relationship between employee perceptions of corporate ethics and organizational culture: An exploratory study. *Asia Pac. Manag. Rev.* **2010**, *15*, 251–260.

86. Visser, M.; Van der Togt, K. Learning in public sector organizations: A theory of action approach. *Public Organ. Rev.* **2016**, *16*, 235–249. [CrossRef]

87. Yeo, R.K. Revisiting the roots of learning organization: A synthesis of the learning organization literature. *Learn. Organ.* **2005**, *12*, 368–382. [CrossRef]

88. Fisher, J.W. Objectivity in psychosocial measurement: What, why, how. *J. Outcome Meas.* **2000**, *4*, 527–563. [PubMed]

89. CNBC. Available online: http://www.cnbc.com (accessed on 22 January 2017).

90. Forbes. Available online: https://www.forbes.com (accessed on 22 January 2017).