Article

A Study of Trust and Cooperation in the Nzoia River Basin Using a Water Policy Game

Abby Muricho Onencan *[iD], Bert Enserink[iD] and Bartel Van de Walle

Policy Analysis Section, Multi-Actor Systems (MAS) Department, Faculty of Technology, Policy and Management, Delft University of Technology, Building 31, Jaffalaan 5, 2628 BX Delft P.O. Box 5015 2600 GA Delft, The Netherlands; B.Enserink@tudelft.nl (B.E.); B.A.vandeWalle@tudelft.nl (B.V.d.W.)
* Correspondence: a.m.onencan@tudelft.nl; Tel.: +31-15-2781-810

check for updates

Abstract: Nzoia river basin county governments barely cooperate in water resources management to jointly increase the basin's food and energy productivity levels, due to limited trust. In this paper, we propose a game-based approach that can be replicated in any river basin, to assess trust and collaboration processes. In particular, we used the pre-game, in-game, and post-game assessment results to assess the relationship between Cooperation and Competition; Trust and Trustworthiness; Trust and Distrust; and (Dis) trust, Complexity, and Uncertainty. The initial assessment of respondents' propensity to trust (PTS) was divided into two variables (trust and trustworthiness) while adopting the unidimensional view of trust and distrust. We later examined whether we could separate the two constructs using a multidimensional scaling (MDS) technique known as the ALSCAL procedure. There are potentially significant results. Namely, that: trustworthiness and trust are not complementary; both cooperation and competition coexisted and increased throughout the game; more profound complexity and uncertainty led to an increment in trust, and reduced complexity and uncertainty led to a decrease in distrust. Based on the results and discussions, we provide recommendations for further research on trust, trustworthiness, and distrust in the river basin management context.

Keywords: trust; trustworthiness; distrust; water cooperation; competition; complexity; deep uncertainty; risk perception; Nzoia river basin; water policy gaming

1. Introduction

Studies show an increasing difficulty for countries to make all their food and energy, within national geographical boundaries, due to scarce water resources [1–10]. Therefore, basin states are faced with a difficult decision: whether to maximize food and energy production or limit and/or stop production and buy the shortfall from other riparian states [11]. Rational decision-making supports cooperation aimed at maximizing production while minimizing costs, regardless of where the food or energy is produced within the basin [6]. Based on previous research, trust is a critical element needed when making the bold decision to stop or limit production and buy the shortfall from other riparian states [9,10,12]. The willingness by the parties to cooperate is sufficient to emit signals that get reciprocated and based on repeated reciprocation of signals that foster cooperation, trust can be cultivated [13] (p. 225). Some riparian states are not willing to cooperate, due to low levels of trust [7,9,10,12,14,15]. Therefore, many basin states are locked in an impasse: they need to cooperate to build trust, and they cannot cooperate because they lack the 'willingness to cooperate,' which is a core element that initiates cooperation [13] (p. 225).

We define a trusting relationship as one where the trustor (A) has definite feelings of assurance and hope that the trustee (B) will act in the trustor's favor "to do X" [16], and not take advantage of

the relationship to the detriment of the trustor [16–18]. According to Hardin [18], the relationship is divided into three parts "A trusts B concerning matters X." In the case of basin states, A can trust B with food production for A's residents, but not energy production. In other instances, A can trust B with money and not personal secrets. Therefore, X is a critical component of the trust relationship, and A can trust B on some issues and not others [16].

One critical advantage of trust is the reduction of complexity [19,20]. In this paper, we define complexity as decision makers inability to assess future effects of planned actions due to unknown interactions between many variables [21]. The past is used to reduce the complexity of the familiar present, through repeated actions reinforced when they lead to the same outcome [22]. However, when the present and future is uncertain, trust is one possible mechanism for reducing complexity. We define uncertainty as the absence of knowledge [23] (p. 16), or when the available knowledge is not certain (impacts of climate change, unknown adaptation and mitigation costs, unknown effects of policy options, and unpredictable social and political environment) [24] (p. 160). According to Luhmann [25] (p. 23), the act of trust reduces deep uncertainties and complexity of the future world. Even though there are many plausible future possibilities, trust reduces the possibility to one possible outcome, the act X by the trustee (B).

Since the fulfillment of X is dependent on B taking into account the interests of A, then B's trustworthiness encapsulates A's interest. Ben-Ner [26] (p. 65) defines trustworthiness as "the willingness of a person B to act favorably towards person A when A has placed an implicit or explicit demand or expectation for action on B." For instance, in a river basin, A (downstream riparian government) entrusts B (upstream riparian government) to maintain good water quality upstream (X). In this instance, it is in the interest of A for B to cooperate and not pollute the river (X). Conversely, it is B's interest to act in a trustworthy manner, and thus maintain good diplomatic relations with A. Therefore, B's trustworthiness encapsulates A's interest.

According to Mayer et al. (1995), trustworthiness consists of three factors: benevolence, ability, and integrity [17]. Evans and Revelle (2008), define benevolence as "the general desire to do good" [27] (p. 1586). Benevolence is the willingness to support others, notwithstanding the costs [27]. Ability refers to a set of competencies, skills, and characteristics that facilitate the effective operation in a certain discipline or domain. Evans and Revelle (2008), define integrity, as "the desire to uphold rules and social norms" [27] (p. 1586). Both benevolent and persons of high integrity reciprocate in a trust relationship [17]. However, the driving force for reciprocity differs. According to Evans and Revelle (2008), the benevolent reciprocate because they are concerned and have a desire to help and perform good actions. Contrary, integrity driven individuals reciprocate out of the conviction that "it is the right thing to do" [27] (p. 1586). If B cannot maintain good water quality (maybe B cannot monitor and stop water polluters), then B will not be considered trustworthy, even if B demonstrates integrity and/or benevolence. Research indicates that evidence of being trusted by the trustor increases the likelihood of the trustee reciprocating [27].

B may choose not to reciprocate and thus take advantage of the trust relationship and act to the detriment of A. The perception by a trustor (A) that the trustee (B) will reciprocate or not, introduces the third element of our study, distrust. According to Gambetta (1988) "it is important to trust, but it may be equally important to be trusted" [13] (p. 221). In trust, the trustor is willing to take the risk of being vulnerable to the trustee, whereas distrust feelings are risk-averse. The distruster is not willing to be vulnerable to the distrusted and take any risks that arise from cooperative actions. Distrust feelings create fear, avoidance, and discomfort. Distrust helps to defend the distruster from the distrusted [28]. In a high trusting relationship, there can also coexist very high distrust levels [28–30]. Lee (2018) explains that distrust has for a long time been perceived as "the opposite of trust." Recent studies have discounted the unidimensional view of trust and distrust and adopted a bi-dimensional view [28,31,32]. Lee (2018) states that "high trust is not the same as low distrust" and argues that distrust is distinct from trust and should be treated as a separate construct [28,31,32]. Thus, low levels

of trust are not similar to high distrust levels, and an increase in trust does not automatically lead to decreased distrust [28,32,33].

Most studies focus on understanding trust [18,25,26,34–38], with insufficient research on trustworthiness and the relationship between trust, trustworthiness, and distrust [34]. Reiersen [34] (p. 1) states that it "is somewhat surprising given that trustworthiness is fundamental to trust." Hardin [16] explains that most of the trust literature barely mentions the term trustworthiness, "though implicitly much of it is primarily about trustworthiness." Reiersen [34] (p. 1) adds that "It makes no sense to trust others if others are untrustworthy." Also, it is detrimental to the trustee to trust an untrustworthy person. According to Reiersen [34] (p. 1) trust is based on the belief that the trustor can be entrusted not to abuse the trust shown. Furthermore, some models theorize that the trust–trustworthiness relationship is complementary and cyclical [27,39]. Moreover, there is limited research on the relationship between trust and distrust, especially in the context of water cooperation [6]. Also, due to weak conception of trust, many studies do not incorporate the three elements of trust in their research instruments: trustee, trustor, and the trustee's behavior (X). Countless researches focus on A and B with no reference to X [40] (p. 19). Hardin [16] explains that A and B without X is an incomplete understanding of trust. Bauer [40] (p. 19) further states that the inclusion of X has diffused slowly within the trust scholars. Bauer [40] (p. 20) recommends that trust scholars should conceptualize trust and formulate questions that demonstrate the three dimensions of trust.

To increase trust amongst riparian governments that share a river basin, it is essential to understand trust elements and how they relate to one another [34] (p. 3). We seek to study trust and collaboration processes using a game environment. We chose policy gaming method because of the low risks associated with gaming, which provide a safe learning and policy practice environment. Through the game, we simulated the river basin context and facilitated the process of interaction, joint problem solving, and learning [41–44]. We used the pre-game, in-game, and post-game assessment results to assess the relationship between:

1. *Cooperation and Competition*: Are cooperation and competition alternatives, or can they coexist [13] (p. 215), in the context of river basin management?
2. *Trust and Trustworthiness*: Is the trust–trustworthiness relationship complementary [27,34,39], within the context of the Nzoia WeShareIt game?
3. *Trust and Distrust*: Is the trust—distrust relationship bi-dimensional [28,31,32], within the context of the Nzoia WeShareIt game?
4. *(Dis) Trust, Complexity, and Uncertainty*: Increased trust leads to the reduction of complexity and uncertainty [25,45]. What is the effect of reduced complexity and uncertainty on distrust [46], within the context of the Nzoia WeShareIt game?

In this research, we undertake a subjective assessment of respondents' propensity to trust (PTS) which we divided into two variables (trust and trustworthiness). We adopt the unidimensional view of trust and distrust at the start of the experiment and later assess whether we can separate the two constructs using the multidimensional scaling (MDS) technique known as ALSCAL procedure [28,33]. The ALSCAL procedure facilitates detailed analysis of the underlying dimensions and clusters using the multidimensional scaling technique. Based on the MDS results, we provide recommendations for further research on trust, trustworthiness, and distrust in the river basin management context.

The paper is structured as follows. Section 2 introduces the Conceptual Framework and explains the research methods and materials which include the experiment design, the materials used in the quasi-experiment, the process of data collection, and the methods used to analyze the data. Section 3 contains the findings based on the in-game data, the initial descriptive statistic results, the Chi-square goodness of fit test results, and the MDS ALSCAL procedure. The subsequent section discusses the findings and explains the limitations of the study. The final section provides concluding remarks and proposes future research.

2. Materials and Methods

We divided the materials and methods section into three parts, first the conceptual design, then the description of the game and finally the methods used to analyze the game data. We used three methods to analyze the pre-game and post-game assessment results. First, the Principal Component Analysis (PCA), followed by the Chi-Square test for goodness of fit, and finally the multidimensional scaling using the ALSCAL procedure. For the **in**-game data, we extracted the results of all the trading rounds and assessed them using Tableau professional edition version 10.2.3.

2.1. The Nzoia WeShareIt Game Conceptual Framework

The Nzoia WeShareIt conceptual framework (Figure 1) was developed to incorporate trust, trustworthiness, and cooperation, in the policy game. It combines three design approaches the: Klein (1993) Recognition-Primed Decisions (RPD) model [47,48], the input-process-output model of serious game design developed by Garris, Ahlers, and Driskell (2002) [49], and Landers (2014) theory of gamified learning model [50]. Onencan (2018) [11] explains how the design approaches were incorporated in the Nzoia WeShareIt game (see supplementary information S1).
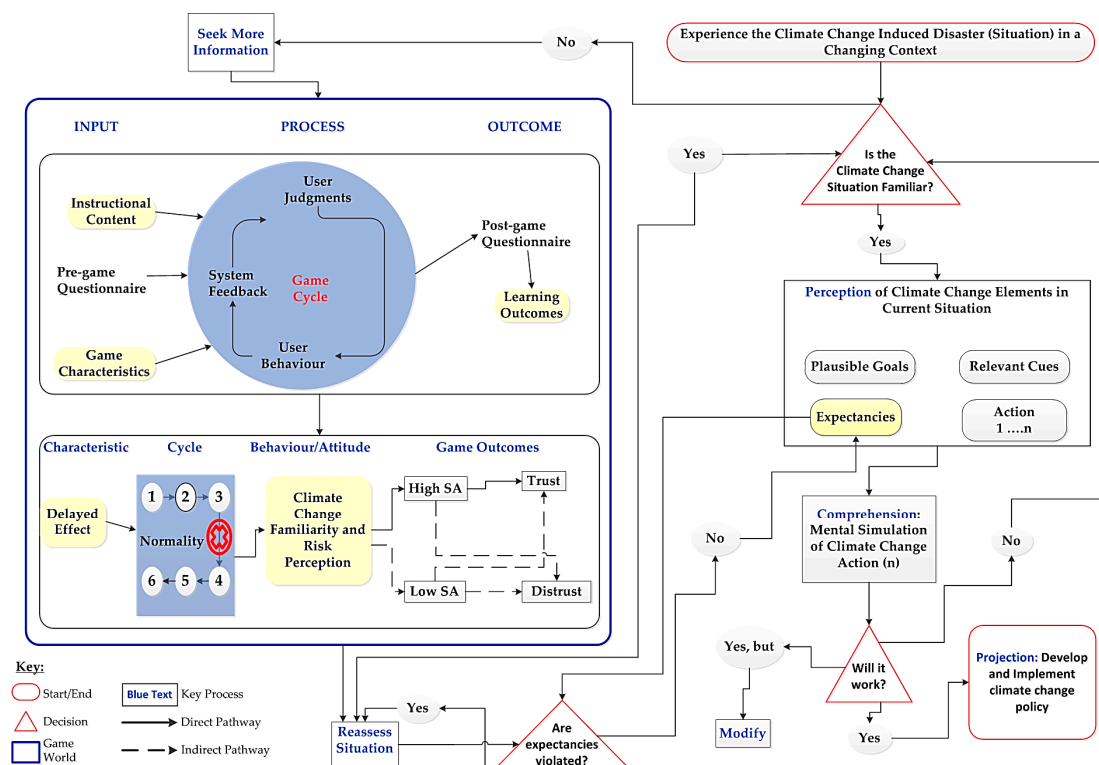

Figure 1. Recognition-Primed Decisions (RPD), Expectancy and Input-process-output model of Nzoia WeShareIt Game. The arrows indicate the causality path. Adapted from Klein [48], Garris, Ahlers, and Driskell [49], and Landers [50]. The game outcomes on situation awareness (SA) are discussed in Onencan [10]. The first three-game cycle rounds represent the status quo (normality). In the fourth round, the player's food and energy resources are halved due to a slow-onset disaster in the form of drought, thus disrupting the status quo.

The Garris, Ahlers, and Driskell (2002), input-process-output model and the Recognition-Primed Decisions (RPD) model [47,48], informed the design of the Nzoia WeShareIt game steps and cycles to ensure structured game design and participant learning. Each cycle provides an opportunity for the players to learn through interactions with other players and Real-time feedback through performance graphs, leaderboard, and in-game reflections [49].

Landers (2014) theory of gamified learning model was adopted to influence behavior and attitudinal change. Since the input-process-output model is designed to influence direct learning, we incorporated the Landers (2014) theory of gamified learning model to influence players behaviors and attitudes, indirectly through the delayed effect game mechanic in the form of climate change-induced disasters (drought) [50]. Drought increases climate change risk perception and is expected to contribute to cooperation ultimately.

2.2. Experimental Design

The game was designed for policymakers in the Nzoia river basin in Western Kenya. The basin comprises six country governments, namely, Busia, Bungoma, Kakamega, Siaya, Trans Nzoia, and Uasin Gishu. Each county government has different resources and different targets. Some counties (Trans Nzoia and Kakamega) have higher food productivity capacity, and others (Uasin Gishu and Bungoma) have higher hydro-electric energy productivity capacity. Some county governments (Busia and Siaya), cannot produce the bare minimum of food and energy for their residents, without compromising economies of scale, the environment, and future generations. Downstream counties (Busia and Siaya) that have lower food and hydro-electric energy productivity capacity could stop food and energy production. Upstream counties could choose to focus on only food or energy production, based on their comparative advantage and stop the production of food (in the case of Uasin Gishu and Bungoma) or energy (in the case of Trans Nzoia and Kakamega).

In the game, riparian county governments are generally faced with three basic policy decision options: Maximize, Limit or Stop food and energy (hydro-electric energy) production. The Maximize decision option is generally based on comparative advantages centered on their food and energy productivity levels and the amount of available water. The Limit decision option requires the riparian governments to minimize their energy and food production especially in the areas where the productivity levels are low. The Stop decision option requires riparian governments to stop energy and/or food production and buy the energy and/or food shortfall from other riparian governments.

The players are united by the shared goal to manage the river basin sustainably. To meet the shared goal, players must communicate their needs and negotiate with county governments to help them meet their needs. Also, players may strategize on winning together. The joint strategy requires a higher level of trust that the other players will keep their promise and not act to their detriment.

We held the game sessions in July 2016. There were seven (7) game sessions played by five (5) policymakers (total of 35 policymakers). Amongst the 35 participants, 12 were female, and 23 were male. The participants were mainly in the 25 to 34 (11 participants); 45 to 54 years (10 participants), and 35 to 44 (7 participants) age ranges. The highest levels of education were a Bachelor's degree (20 participants) and a college diploma (7 participants). The sessions were conducted first in Busia, then Kakamega, followed by Bungoma and finally Trans Nzoia county government. Onencan et al. (2018) provide a detailed description of the game design, participants profiles and the sessions [11].

2.3. Materials

The assessment contained 18 questions from the Propensity to Trust (PTS) scale. The PTS scale is context-specific, it measures the underlying behavior of a person based on the simulated state [17,27,51]. Ten of the questions assess trust and eight assess trustworthiness. For each question, the respondent assessed the accuracy of the statement, according to their perception. The scores were from 1 to 5, one meaning "very inaccurate" and five meaning "very accurate."

The original PTS scale has 21 questions. We adopted the PTS scale from Evans and Revelle (2008), with a few modifications. Appendix A encompasses a description of the modifications made and Table A1 in Appendix B lists all the questions asked in the pre-game and post-game questionnaire. Supplementary Materials S2, contains the raw data we used to extract the results.

2.4. The Methods

2.4.1. Method 1: Principal Component Analyses (PCA)

Using the respondent's responses, we first conducted a PCA of the two sub-scales to assess the underlying structure of the two PTS sub-scales psychometrically. The purpose was to check whether the two sub-scales had sufficient loadings before undertaking further analyses [27]. A detailed explanation of the PCA results is in Appendix B.

The trustworthy factor loadings were stable and significantly higher than the trust factor loadings in both the pre-game and post-game stages. Therefore, we concluded that the trustworthy sub-scale measures one construct.

The trust sub-scale was very unstable and seemed to be measuring more than one construct. Though the trust sub-scale passed the goodness of fit test at the post-game level, its loadings are not high, and the underlying structure was not straightforward. Thus there was the need for further analyses, as explained in the subsequent sub-sections. The loadings per item for the two subscales are in Appendix B (Table A1).

2.4.2. Method 2: Chi-Square Test for Goodness-of-Fit for PTS

The Chi-Square goodness-of-fit test was used to compare an observed distribution with a theoretical distribution. The Chi-Square test for goodness-of-fit null hypothesis is: there is no significant difference between the post-game results for the PT sub-scales (trust and trustworthiness), and the pre-game results (H0). The Chi-Square goodness-of-fit test alternative hypothesis is: there is a significant difference between the post-game results for the PT sub-scales (trust and trustworthiness), and the pre-game results (H1). We reject the null hypothesis if Sig. < 0.05.

2.4.3. Method 3: The Multidimensional Scaling (MDS) ALSCAL Procedure

Multidimensional scaling (MDS) is a technique used to visualize the level of similarity of the individual objects in a dataset [52]. It places these objects in an n-dimensional space, the coordinates of which are formed by a series of hidden or underlying attributes [53]. The purpose of MDS is to identify those attributes, compute the coordinates of each object and represent the objects in space [52]. The primary purpose of conducting MDS was to compute the distances between objects and group them in clusters, based on their similarities [52]. Afterward, we labeled each dimension according to the characteristics of the object in each class. The MDS procedure starts from a single object attribute to discover the underlying dimensions behind that attribute [53].

3. Results

In this section, we present the findings on the WeShareIt Game trade exchanges (3.1), the PT sub-scales responses findings based on the Chi-Square goodness-of-fit test (3.2) and the MDS (3.3). Appendix C explains the results of the pre-game and post-game descriptive statistics. The standard deviations and means for each PTS sub-scale item, are contained in Table A2.

3.1. Cooperation Outcomes: Nzoia WeShareIt Game Results

Using Tableau, we visualized all the trades conducted by the five county governments (Figure 2). The exchanged resources are food, energy (hydroelectric power), and money [11]. In-game trade data was extracted to visualize the exchanges between County Governments starting with hydro-electric energy, then food, and finally money. There were numerous trade transactions; some can be grouped as 'short-term opportunistic' transactions, that did not develop long-term engagement while others were grouped under strong long-term oriented relationships [54] (p. 338).

However, the short-term opportunistic transactions are minimal compared to the long-term oriented and repeated transactions. Due to the repeated transactions, a network of buyers and

sellers seemed to be emerging based on comparative advantages. Uasin Gishu is the sole provider of hydro-electricity energy for the basin. Trans Nzoia and Kakamega are the food providers. However, Trans Nzoia provides more food than Kakamega. The primary consumers that ensure that the excess food and energy are utilized are Bungoma and Busia.
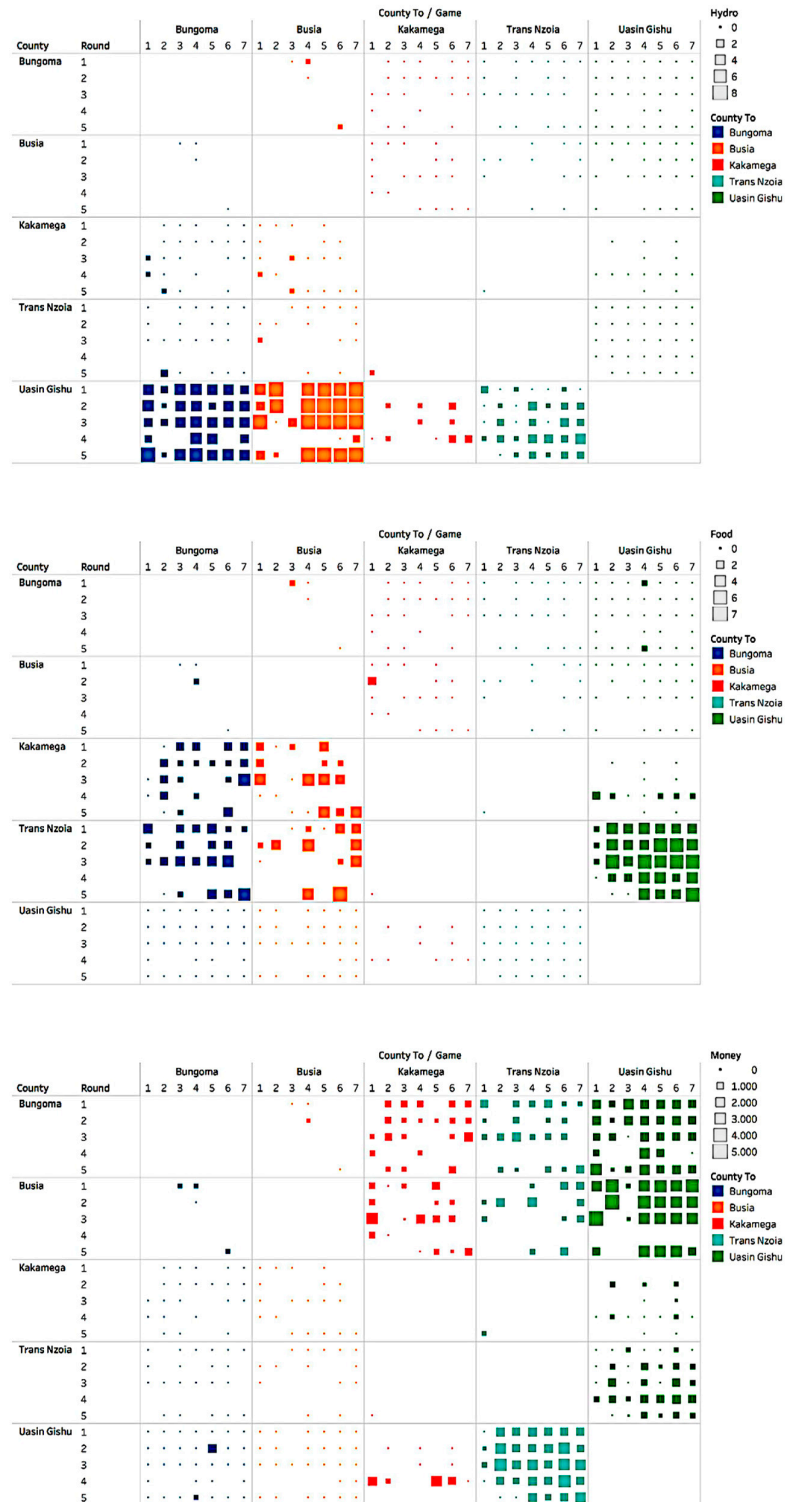

Figure 2. Exchanges between County Governments: (**a**) Hydro-electric energy, (**b**) food, and (**c**) money.

3.2. Chi-Square Test for Goodness-of-Fit for PTS

The goodness-of-fit results indicate statistical differences in the trustworthy sub-scale for all the eight variables. Based upon the observed frequencies it appears that there was a significant increase or decrease in trustworthiness between the pre-game and post-game results. The *p*-value for all the eight variables is lower than 5% (Sig. < 0.05). Hence, we reject the null hypothesis: there is no significant difference between the post-game results for the trustworthiness sub-scale, and the pre-game results (H0). Hence, there is an enormous variance between the post-game results for the trustworthiness sub-scale, and the pre-game results (H1).

The results for the trust sub-scale are varied. Based upon the observed frequencies, it appears that there was a significant increase or decrease in trust between the pre-game and post-game results in six variables (T1, T2, T3, T4, T7, and T10). The *p*-value for all the six variables is lower than 5% (Sig. < 0.05). Consequently, we reject the null hypothesis: there is no significant difference between the post-game results for the trust sub-scale, and the pre-game results (H0). As a result, there is a significant difference between the post-game results for the trust sub-scale, and the pre-game results (H1). Additionally, there was no significant increase or decrease in trust between the pre-game and post-game results in four variables (T5, T6, T8, and T9). The *p*-value for all the four variables is higher than 5% (Sig. < 0.05). Therefore, we maintain the null hypothesis: there is no significant difference between the post-game results for the trust sub-scale, and the pre-game results (H0). The results of the trustworthy sub-scale Chi-Square test for Goodness-of-fit are in Table A3 and for the trust sub-scale in Table A4 (Appendix D).

3.3. Multi-Dimensional Scaling (MSD) Using the ALSCAL Procedure

To be able to assess the multiple dimensions of the PTS, we performed a multidimensional scaling (MDS) using the ALSCAL procedure. Appendix E describes the ALSCAL procedure. To assess the consistency of the model we considered the scatterplot of linear fit (Figure 3). Since the points in the chart tend to gather around the chart diagonal (straight line), the model's consistency is confirmed.
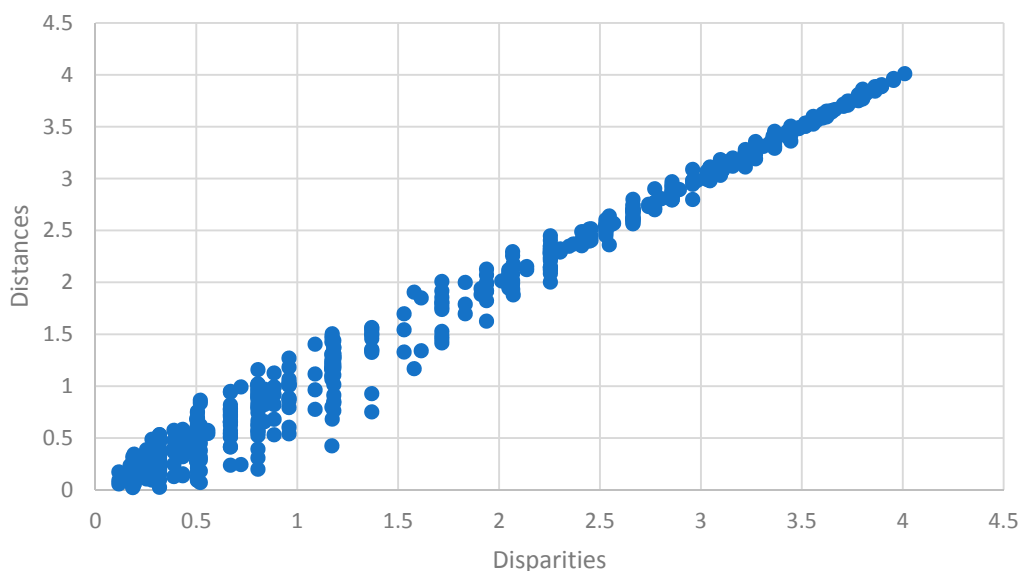

Figure 3. Scatterplot of linear fit from the Euclidean, distance model.

Based on the scatterplot of linear fit that was derived from the Euclidean distance model, there was more extensive space for disparities than distances. The disparities ranged from slightly above 0 and 4. However, the distances were narrower than the disparities because the responses seemed to gather around the chart diagonal. As the disparities increased, the distances decreased. There were more considerable distances on the lower side of the Y-axis where the disparities were less compared

to the upper side of the Y-axis. Figure 4 visualizes the distances and disparities separately based on the 36 observations/stimuli.

After that, we assessed the stimulus coordinates, configured in a two-dimensional space. There were 36 observations or stimuli (18 results from the pre-game questionnaire on trust and 18 results from the post-game questionnaire on trust). For each of the 36 observations, the model computed two coordinates (Dimensions 1 and 2) and grouped the 36 observations into clusters. Figure 5 plots the two-dimensional solution obtained for the pre-game and post-game individual dissimilarity scores for the trust and trustworthiness sub-scales, grouped in two dimensions.
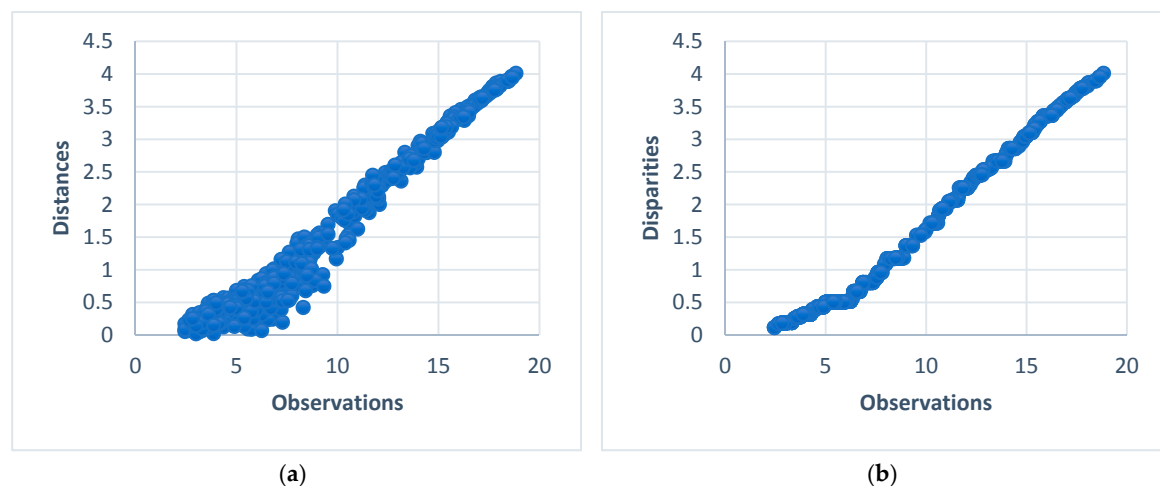

(**a**)　　　　　　　　　　　　　　　　　　　(**b**)

Figure 4. Scatterplots based on the 36 observations/stimuli: (**a**) Scatterplot of Nonlinear Fit from the Euclidean, distance model; (**b**) Transformation scatterplot from the Euclidean, distance model.
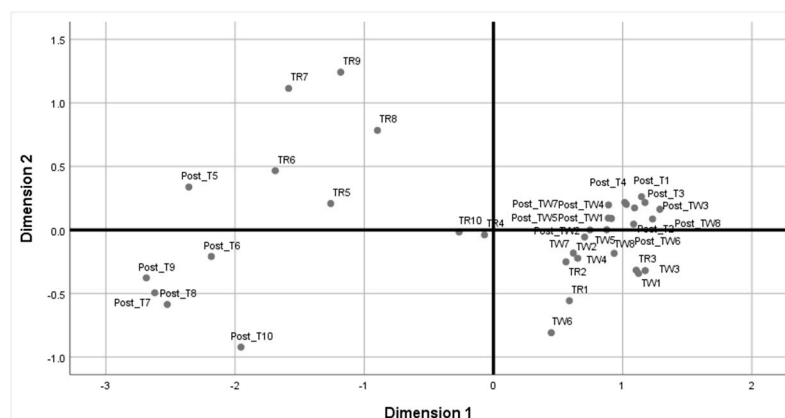

Figure 5. Conceptual Map of the Derived Stimulus Configuration from the Euclidean, distance model.

We clustered the conceptual map results into four groups. Figure 6 and Table A5 (Appendix E) contains the clustering of the 36 trust and trustworthy observations based on the derived stimulus configuration from the Euclidean distance model into four clusters, under two dimensions. To understand the components of the four clusters, we first assessed the features of each of the 18 variables within the PTS. All the 12 variables in the first cluster assessed the respondent's positive perceptions, beliefs, and actions (trust). The four clusters mainly contain the results of the 12 positive, trustworthy PTS sub-scale in one cluster and the results of the six negative trust sub-scale in another. Based on the respondent's score, high scores indicated a high disposition to trust. Low scores indicated a low disposition to trust. The last six variables were grouped in the second cluster that assessed the respondent's negative perceptions, beliefs, and actions (distrust). Based on the respondent's score, high scores indicated profound distrust and low scores indicated low distrust. The other two clusters

differentiate the pre-game results from the post-game results. Particularly, two variables did not fit into the four clusters: T10 and T4.

After that, we interpreted the two dimensions and gave suggestive labels for each dimension: (1) Dimension 1 represents uncertainty, and (2) Dimension 2 represents complexity (Figure 6). The first dimension refers to the level of uncertainty with two values—low and high—and the second dimension denotes to the level of complexity with two values—low and high. We were guided by the elements of trust as highlighted by Gambetta [13] (p. 218) when determining the labels for the two dimensions. The first element relates to trust being a "threshold point, located in a probabilistic distribution." In this threshold, the values range from 0 to 1, with 0 representing complete distrust and 1 representing complete trust. At the midpoint of these values (0.50) is uncertainty. From the graph, the midpoint of trust and distrust is located along Dimension 1, justifying our interpretation that Dimension 1 represents uncertainty. Secondly, trust is principally relevant when there is deep uncertainty and high complexity. Luhmann [25] explains that a critical element of trust is the reduction of complexity. Thus, Dimension 2 represents complexity.
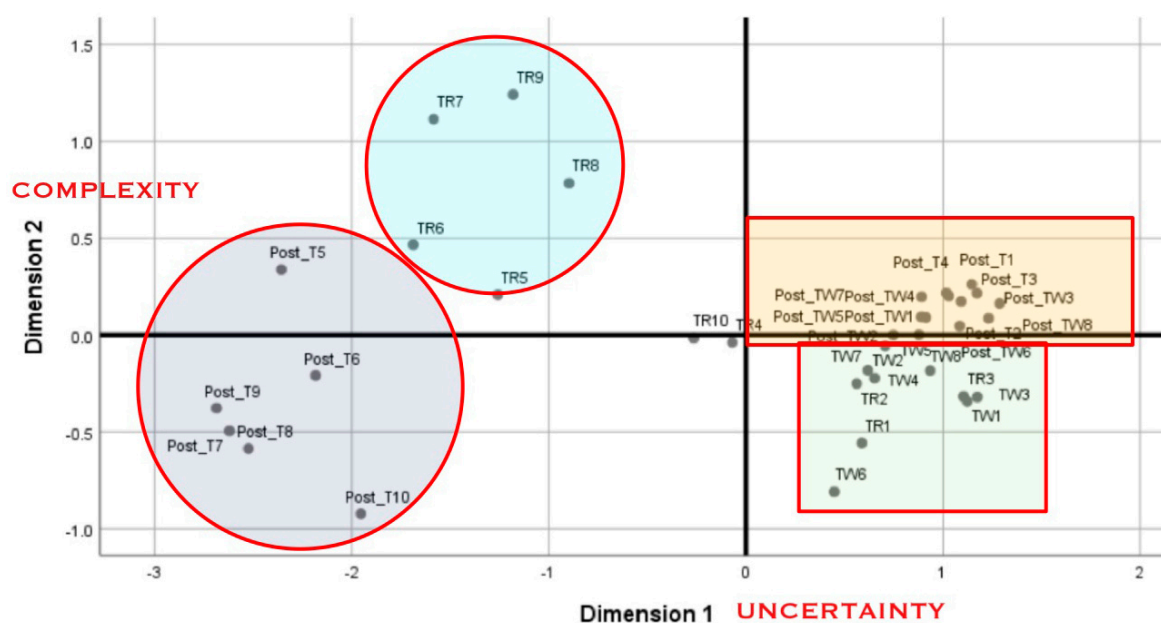

Figure 6. Clustering of the Conceptual Map based on the two dimensions. Dimension 1 represents the impact of high and low uncertainty on the levels of trust and distrust, with the positive impacts on the right side and negative impacts on the left side. Uncertainty had a mixed marginal impact on trust and a distinct negative impact on distrust. Dimension 2 represents the impact of high and low complexity on the levels of trust and distrust, with trust increasing and distrust decreasing.

4. Discussion

The first four parts of the discussion are centered on the four research questions, as highlighted in the Introduction section. The fifth part discusses the limitations of this particular study.

4.1. Cooperation and Competition

Cooperation and competition increased throughout the game. The results support the findings by Gambetta [13] (p. 215) that competition and cooperation should not be perceived as alternatives because they coexist throughout the game sessions. Water cooperation was evident in the game, and at the same time, the participants chose not to stop unproductive food and energy production, as a safety net. None of the county governments decided to make 100% of their food and energy and not engage in any trade relations with the neighboring counties. Based on the debriefing sessions, the respondents stated that it was not wise to completely trust the neighboring counties to produce their food or energy.

Therefore the small production, though under unproductive conditions, was a safety net, intended to protect them from the 'increased vulnerability' that emerges from a trust relationship [54] (p. 338). Competition and cooperation coexisted due to decisions not to stop unproductive food and energy production. The primary challenge in the game was finding a healthy balance between cooperation and competition [13] (p. 214).

4.2. Trust and Trustworthiness

Apart from the eight trustworthy variables, there were 12 trust variables. Within the 12 trust variables, four assessed the respondent and six assessed other players in the game. Therefore, there were three groups of constructs being measured: trustworthiness (8 variables TW 1-TW8), cooperative nature of the respondent (4 variables—T1–T4) trust (6 variables T5–T10). All these 18 variables comprise the PTS/DTS.

The game had a significant positive impact on individual perceptions of their trustworthiness. The players exhibited high trustworthiness, and high ability to cooperate. All the respondents had positive perceptions of their trustworthiness and this perception increased significantly after the game sessions.

The players' assessments indicated a low perception of the trustworthiness of other players. Trust levels declined consistently after the game sessions. The trustee (B) was reciprocating (based on the in-game results) with clear evidence of distrust by the trustor (A).

The trust results were not complementary to trustworthy because the ratings of B by A worsened after playing the game (except the self-assessment PTS ratings, namely T1–T4). In the case of the Nzoia WeShareIt game, the trust-trustworthiness relationship was not complimentary. The initial Chi-Square test for goodness-of-fit test results indicates contrary findings to previous research regarding the complementarity of the relationship between trustworthiness and trust [27,34,39].

According to Reiersen [34] (p. 4) the trust problem is the lack of knowledge by the trustor that the trustee is trustworthy and B is the solution. To overcome the trust problem, B has to convince A, that B is trustworthy. The 35 Bs, in the case of the Nzoia WeShareIt game, did not manage to convince the 35 As, that they are trustworthy. The consistent poor rating of other players indicated that the interactions between the players led to conflicts over the shared water resources, that could not be retracted, leading to low B trustworthiness ratings [28,55]. Increased competition, exclusive dealing, price fixing, refusal to trade with others, and absorption of a competitor led to unfair trading practices. Though cooperation improved, there was also augmented competition, at the expense of B's trustworthiness.

There are various explanations of these contrary findings. First, in the initial game rounds, cooperation was not based on trust, leading to conflicting results. Disposition to trust was consistently built throughout the game because there were repeated actions that proved that cooperation works. Reierson [34], explains that cooperation based on repeated interaction "is just pure calculation and maximization of long-term self-interest." Therefore, the initial cooperative actions were not grounded on trust and this may be one of the reasons for the contradictory results [34], Trust is one of the elements that enable cooperation, and we could not assume that because there were cooperative actions, trust was established at the initial stages of the game.

Second, Ben-Ner and Halldorsson [26] explain that trustworthiness is embedded in norms. Reiersen (2018) adds that "trustors trust because they are aware that trustworthiness is rooted in norms." Hardin [18] (pp. 25–29) explains some difficulties experienced when seeking to measure trustworthiness in various societies. The first difficulty is framing of the question: different societies frame the same question differently leading to varied responses. The different frames are mainly due to different institutional structures or variation in time. Another difficulty is individuals with a background of untrustworthy relations. It is harder for such individuals to undertake the risk of perceiving someone else as trustworthy and thus they cannot easily develop a trusting relationship.

Third, changes in the game led to instances of incentive incompatibility that reduced B's level of trustworthiness. If B is a producer of food or energy, X refers to four things:

1.　B will make the relevant water allocation to produce excess energy and/or food on behalf of A;
2.　B will sell the excess production to A before considering other competing buyers;
3.　B will not insist on high fixed prices, even when there is high demand; and
4.　B will strive to keep the promise made to A, even under unforeseen circumstances (droughts).

Based on the game design, there is a high likelihood that B will not fulfill all the Xs in every round. In the drought round, B loses half of their resources, and even though they fulfilled the first X, they cannot fulfill the last X because the game rules require B first to ensure that its residents have sufficient food and energy before selling to others. Therefore, though B's actions are perceived untrustworthy, it was challenging for B to convince A that they are trustworthy when achieving specific game rules competed with an act of trustworthiness. The drought round indicates that there are times when B's interest does not encapsulate A's [16,18] thus, B had no interest in being trustworthy. Consequently, the post-game questionnaire assesses a summation of all the interactions and does not take account of the moments, when B's interest did not encapsulate A. In these instances, trust and trustworthiness were not complimentary. Incentive incompatibility and competition led to consistent low ratings for B's trustworthiness. Thus confirming the statement by Hardin [18] that trust is "heavily limited and conditional." X is a critical component of trust and B's interest must encapsulate A, before a complementary and cyclical relationship is established.

4.3. Trust and Distrust

The trust–distrust relationship is bi-dimensional [28,31,32], within the context of the Nzoia WeShareIt game. The results imply that a single dimension analysis of trust of the overall level of PTS may be potentially misleading because the respondents were measuring other constructs, not envisaged when developing the assessment tool. The distances between the variables indicate that the respondents measured multiple constructs. Disposition to trust variables all clustered into one big group, while another group clustered on the left side of the graph. The vast distances between these two groups confirm that the multidimensional scale separated the two constructs. Though the trustworthy sub-scale was not affected, the trust sub-scale measured different constructs leading to conflicting results.

The results confirm that trust and distrust are distinct concepts. Some constructs were being measured, namely, trust in others, distrust of others, a sense of trustworthiness of others, and a self-assessment of trustworthiness. There were high scores for self-assessment of trustworthiness and distrust of others. The respondents were facing a difficult time being trusted by other players. Each player considered him/herself highly trustworthy, but that perception was not mutual neither was it a collective perception. The general perception was that none of the players could be trusted. Based upon the vast distances between all the distrust results on the conceptual map, there was no significant change in distrust levels at the start and the end of the game for all the six variables.

4.4. Trust, Distrust, Complexity, and Uncertainty

Increased complexity and uncertainty led to increased trust. Based on Figure 6, trust is located in the area in the graph representing deep uncertainty and high complexity. The research results indicate that trust levels increased as uncertainty and complexity increased.

Moreover, trust is excluded in circumstances where a trust relationship will not affect the decision made. In the graph, trust is excluded from the left side of the graph, where there is low uncertainty and low complexity because in this case trust is not needed to make a decision [13,25,56]. It is also excluded from the upper left of the graph where there is low uncertainty and high complexity, where efforts to understand and address the complexity are needed, instead of trust. Klinke and Renn [23] (p. 1086) support this finding by stating that "it does not make sense to incorporate . . . perceptions,

or any other social aspects into the function of resolving (cognitive) complexity. Charnley [57] adds that when there is deep complexity, social solutions can be used if the specific community has certain knowledge that is critical for reducing complexity. Therefore, community knowledge replaces trust thus reducing complexity, where the situation is certain and unambiguous. Klinke [23] proposes technical solutions in the form of cost-effective methods to address deeply complex problems that are unambiguous and certain.

Reduced complexity and uncertainty led to a decrease in distrust. The conceptual map indicates a decrease in distrust between the pre-game and post-game findings, though the decrease was not significant. The decrease occurred when complexity and uncertainty were reduced. Therefore, under familiar conditions (when the players were more familiar with the risk and the other respondents), complexity and uncertainty are reduced leading to a decrease in distrust.

4.5. Limitations of the Study and Future Research

The research study approach faced three main limitations. First, it was not possible to include the X in the pre-game questionnaires, which led to the assessment of generalized or social trust, at the start of the game. Hardin defines general trust as "trust in random others or social institutions without grounding in specific prior or subsequent relationships with others [18] (p. 23)." However, based on the nature of the questions (apart from T5–T10 that needed reverse coding), the response, whether in the real world or a game environment, would not make a huge difference since this was one's subjective assessment of their level of trust and trustworthiness. Therefore the comparison we made between the pre-game and post-game questionnaires is justified because the questions were a subjective assessment of the trustor (A), save the six questions that focused on the trustee (B). However, the effects of this limitation were not considered immense, since there was no significant difference between the distrust scores before and after the game.

Second, the post-game questionnaire was designed to assess a summation of all the interactions. Though this is a convenient and less cumbersome approach, it did not take account of the moments, when B's interest did not encapsulate A.

Third, we adopted both the 'belief-based [58]' and the 'behavioral-based [59,60]' approaches to trust without integrating the two in one assessment tool. Bauer [40] (p. 21) points out that one of the biggest challenges of measuring trustworthiness is the inability to yield credible answers. To address this, we used in-game data to measure trustworthiness actions and asked indirect behavioral questions using the pre- and post-game questionnaires to measure attributes that influence trustworthiness. It was not clear from the in-game data whether the cooperative actions were grounded purely on trust or there was another factor that contributed to the cooperative actions.

5. Conclusions

Trust is a critical element that enhances water cooperation within a shared river basin. Trust establishes a healthy balance between cooperation and competition with the aim of increasing the benefits that riparian governments can derive from the shared water resource. To be able to tap into the positive aspects of trust, it is imperative that trust concepts and collaborative processes that relate to water resources management are better understood.

Little is known of the relationship between trust and trustworthiness. There is limited literature on trustworthiness and how it relates and contributes to a trusting relationship. Moreover, the belief that distrust and trust are unidimensional has led to limited research on the relationship between trust and distrust. Additionally, there is established literature that affirms that trust reduces complexity by limiting the number of possible future options to one: action X by B. However, very little is known regarding the impact of changing levels of uncertainty and complexity on the levels of trust and distrust.

In this paper, we offer a game-based approach that can be replicated in any river basin, to assess trust and collaboration processes. The model is supported by pre-game, in-game, and post-game data

that measure both the 'belief-based' and the 'behavioral-based' approaches to trust. After that, the paper discusses some interesting and potentially important results, namely, that:

1.   Cooperation and competition coexisted and increased throughout the game;
2.   In the Nzoia WeShareIt Game context, trustworthiness, and trust were not complementary;
3.   Trust and distrust are bi-dimensional and operated simultaneously in the game due to the multiple and mixed conditions, leading to varied complexities and uncertainties; and
4.   Increased complexity and uncertainty led to increased trust whereas decreased complexity and uncertainty led to a decrease in distrust.

We recommend that future researches focus on deepening understanding of the relationship between trust, trustworthiness, and distrust in the context of river basin management. Also, there is a need to improve existing measurements of trust so that trustworthiness and distrust are correctly measured. These studies should aim at increasing water cooperation within a shared basin and contribute to complexity and deep uncertainty studies. Finally, there is a need for more research on how to establish a healthy balance between cooperation and competition, once the attributes of (dis)trust are better understood.

Appendix A. Modifications Made to the PTS Scale

First, the original PTS had 21 questions [27]. We deleted three PTS questions, namely:

1.   Stick to the rules (No. 6);
2.   "Return extra change when a cashier makes a mistake [61]" (No. 10);
3.   Retreat from others (No. 15).

We deleted question 6 because the players easily confused it with question 7 (believe the laws should be strictly enforced). One was measuring the *act* of sticking to the rules (No. 6) and the other the *belief* that laws should be strictly enforced. The player might have the belief but not stick to the rules. However, we decided, based on the responses received when testing the questionnaire, that the construct should be deleted to reduce confusion. We also deleted question 15 because it was similar to question 18 (avoid contact with others) and would lead to confusion when the players were providing their responses.

Additionally, we deleted question 10 because in the game there was no cashier and change. Also, money transfers between the players were electronically calculated with a very low probability of cheating or avoidance to make payments. We made deductions during the particular round, in predetermined stages. For instance, we deducted the payment of a penalty at the close of the trading round. Therefore, the players had no chance to evade payment or pay less than what was required.

After adopting the remaining 18 questions, the language was modified to reflect a policy game context. For instance, question 19: we transformed it to believe that most people would lie to get ahead so that it was relevant to the game. The revised question read: believe that most people (players) would lie to get ahead (T8). Another example is: would never cheat on my taxes. We transformed it to never cheated (TW6).

Appendix B. Principal Component Analysis (PCA) Results

We conducted a PCA using XLSTAT. The PCA results are in Table A1.

Table A1. A table containing the PTS items for the pre-game and post-game questionnaires. Changes made to the original questions are in brackets. A principal component analysis was undertaken to assess the underlying PTS structure. The table also contains factor loading of PTS items for both the pre-game and post-game results. The trustworthy sub-scale has high loadings. There were low loadings for the first four trust items at the pre-game stage and high at the post-game stage, though negative.

Questions/Sub-Scale/Factor Loading of PTS Items	Sub-Scale	Pre-Game		Post-Game	
N = 35		TW 37%	Trust 27%	TW 72%	Trust 46%
1. Listen(ed) to my conscience (TW1)	Trustworthy	0.63		0.89	
2. Anticipate(d) the needs of others (TW2)	Trustworthy	0.42		0.72	
3. Respect(ed) others (TW3)	Trustworthy	0.44		0.84	
4. Gets (got) along with most people (T1)	Trust		0.14		−0.73
5. Always been (Was) completely fair to others (TW4)	Trustworthy	0.60		0.86	
6. Believe that laws (game rules) should be strictly enforced (TW5)	Trustworthy	0.63		0.82	
7. Have (had) a good word for everyone (T2)	Trust		-		−0.67
8. Value(d) cooperation over competition (T3)	Trust		−0.28		−0.60
9. Would never cheat on my taxes (never cheated) (TW6)	Trustworthy	0.41		0.84	
10. Follow(ed) through with my plans (TW7)	Trustworthy	0.84		0.89	
11. Believe(d) that people (players) are (were) basically moral (T4)	Trust		0.34		−0.62
12. Finish(ed) what I start (ed) (TW8)	Trustworthy	0.74		0.91	
13. Filled with doubts about things (was filled with doubt) (T5)	Trust		0.78		0.49
14. Feel short-changed in life (Felt short-changed) (T6)	Trust		0.75		0.56
15. Avoid contact with other(s) (players) (T7)	Trust		0.71		0.83
16. Believe that most people (players) would lie to get ahead (T8)	Trust		0.58		0.77
17. Find it hard to forgive others (players) (T9)	Trust		0.61		0.82
18. Believe that people (other players) seldom tell you the whole story (T10)	Trust		0.32		0.64
Valid N (listwise)	32				

Based on the findings, three items were identified not to have sufficient loadings (higher than 0.30), at the pre-game stage, namely:

1. Gets (got) along with most people (T1), factor loading of 0.14
2. Have (had) a good word for everyone (T2)
3. Value(d) cooperation over the competition (T3), factor loading of −0.28

We did not discard these three items because they had very high factor loadings (above 0.60), at the post-game stage. However, these factor loadings were negative leading to the weakening of the VSS. We did not discard any item because the scale and its sub-scales had already been tested and approved by Evans and Revelle (2008), as the goodness of fit for the VSS. However, we noted that Items T1, T2, T3, and T4 might not contribute high loadings for the trust sub-scale.

Appendix C. Pre-Game and Post-Game Trustworthy and Trust Descriptive Statistics

For each of the PTS sub-scales, we computed a two-way contingency table, that cross-classifies the PTS subjective rating before the game and after the game session. The pre-game variables were assigned two different labels: T for trust and TW for trustworthiness. Subsequently, we numbered each variable. For trust, the variables ranged from T1 to T10, and for trustworthiness, the variables ranged from TW1 to TW8. We assigned the numbers based on the numbering in the online questionnaire.

For the post-game labels, we used the same labels, and added the word post, to differentiate the findings from the pre-game findings.

Table A2. Pre-game and post-game descriptive statistics.

Pre-Game and Post-Game Descriptive Statistics		Pre-Game		Post-Game	
		Mean	Std. Deviation	Mean	Std. Deviation
TRUSTWORTHY SUB-SCALE	1. Listen(ed) to my conscience	4.8	0.4	4.5	0.8
	2. Anticipate(d) the needs of others	4.4	0.8	4.4	0.7
	3. Respect(ed) others	4.8	0.4	4.7	0.8
	4. Have always been (Was) completely fair to others (players)	4.3	0.6	4.4	0.8
	5. Believe that laws (game rules) should be strictly enforced	4.5	0.9	4.4	0.9
	6. Would never cheat on my taxes (never cheated)	4.2	1.1	4.6	0.9
	7. Follow(ed) through with my plans	4.4	0.6	4.5	0.8
	8. Finish(ed) what I start (ed)	4.6	0.6	4.7	0.8
TRUST SUB-SCALE	1. Get (Got) along with most people (players)	4.3	0.95	4.6	0.8
	2. Have a good word for everyone	4.1	1.0	4.6	0.8
	3. Value cooperation over competition	4.8	0.4	4.7	0.8
	4. Believe that people are moral	3.7	0.97	4.5	0.8
	5. Filled with doubts about things (Was filled with doubt)	2.9	1.2	2.1	1.3
	6. Feel short-changed in life (Felt short-changed)	2.6	1.3	2.3	1.5
	7. Avoid contact with others (other players)	2.3	1.3	1.9	1.4
	8. Believe that most people would lie to get ahead Post-game—Believed that most players lied to get ahead	3.1	1.3	2.1	1.5
	9. Find it hard to forgive others (found, other players)	2.5	1.4	1.9	1.4
	10. Believe that people (Players) seldom tell you the whole story	3.5	0.8	2.5	1.5

From each of the contingency tables, we computed the mean and standard deviation scores to assess the difference between the pre-game scores and the post-game scores, for each variable (dissimilarity matrix). The difference between the mean scores indicated whether there was positive change (increase in the mean score), a negative change (decrease in the mean score) or no change at all.

Based on the findings, there was an increase in the PT in four variables within the trustworthiness subscale (TW4, TW6, TW7, TW8). The increase in TW4, TW7, and TW8 was marginal (0.1). TW6 was reported to have the highest increase (0.4). There was no change in PT in TW2. However, there was a slight decline in the standard deviation for TW2 (0.1). There was also a decline in PT for three variables (TW1, TW3, and TW5). The decline was marginal for TW3 and YW5 (0.1). The decline in PT for TW1 was the highest (0.3).

According to the initial findings, there was an increase in the PT in three variables within the trust subscale (T1, T2, and T4). None of the increases were marginal (0.3, 0.5 and 0.8). T4 was reported to have the highest increase (0.8). The standard deviations for the trust subscale were much higher than the trustworthiness subscale. The highest standard deviation for trust was 1.5, and for the trust, it was 1.1. There was a decline in PT for seven variables (T3, T5–T10). The decline was marginal for T3 only. The decline in PT for T8 and T10 were the highest (1.0).

Appendix D. Chi-Square Test for Goodness-of-Fit for PTS Results

The use of raw cell frequency to assess whether there was trust formation or not, can be misleading. From the raw cell frequencies (expressed as mean and standard deviation), T8 had the highest mean difference between the pre-game and post-game results (1.0) and standard deviation (1.5). In the non-parametric Chi-Square test for goodness-of-fit for the trust sub-scale, the results for T8 were not significant. Additionally, the trustworthy mean differences between the pre-game and post-game results were marginal compared to the trust sub-scale with much lower standard deviations. On the contrary, all the trustworthy Chi-Square tests for goodness-of-fit results were significant.

Table A3. Pre-game and post-game trustworthy Chi-Square test statistics.

		TW1	TW2	TW3	TW4	TW5	TW6	TW7	TW8
PRE-GAME	Chi-Square	11.765 [a]	25.000 [b]	12.600 [c]	13.086 [d]	29.800 [b]	21.114 [b]	12.057 [d]	17.200 [d]
	df	1	3	1	2	3	3	2	2
	Asymp. Sig.	0.001	0.000	0.000	0.001	0.000	0.000	0.002	0.000
POST-GAME	Chi-Square	16.171 [a]	24.314 [b]	63.171 [b]	14.629 [a]	24.588 [c]	43.971 [b]	15.314 [a]	31.771 [a]
	df	2	3	3	2	3	3	2	2
	Asymp. Sig.	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000

PRE-GAME	a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 17.0. b. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 8.8. c. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 17.5. d. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 11.7.
POST-GAME	a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 11.7. b. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 8.8. c. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 8.5.

Table A4. Pre-Game and Post-Game Trust Chi-Square Test Statistics.

		T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
PRE-GAME	Chi-Square	32.000 [a]	25.143 [a]	12.600 [b]	27.714 [a]	6.571 [a]	5.714 [a]	11.588 [c]	3.714 [a]	7.143 [a]	17.059 [d]
	df	4	4	1	4	4	4	4	4	4	3
	Asymp. Sig.	0.000	0.000	0.000	0.000	0.160	0.222	0.021	0.446	0.129	0.001
POST-GAME	Chi-Square	23.029 [a]	20.800 [a]	57.686 [b]	16.294 [c]	18.647 [d]	16.000 [e]	36.857 [e]	29.879 [f]	23.647 [g]	10.571 [e]
	df	2	2	3	2	4	4	4	4	3	4
	Asymp. Sig.	0.000	0.000	0.000	0.000	0.001	0.003	0.000	0.000	0.000	0.032

PRE-GAME	a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 7.0. b. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 17.5. c. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 6.8. d. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 8.5.
POST-GAME	a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 11.7. b. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 8.8. c. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 11.3. d. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 6.8. e. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 7.0. f. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 6.6. g. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 8.5.

Raw-cell frequency used to assess whether there was trust formation or not, can be misleading, as explained in Appendix D. From the raw cell frequencies (expressed as mean and standard deviation), T8 had the highest mean difference between the pre-game and post-game results (1.0) and standard deviation (1.5). In the non-parametric Chi-Square test for goodness-of-fit for the trust sub-scale, the results for T8 were not significant. Additionally, the trustworthy mean differences between the pre-game and post-game results were marginal compared to the trust sub-scale with much lower standard deviations. On the contrary, all the trustworthy Chi-Square tests for goodness-of-fit results were significant. It appears that there are hidden attributes in the data that cannot be visualized with the use of raw cell frequencies, which emphasizes the need for further investigation. Also, based on the high standard deviations in the trust sub-scale, the respondents seem to measure more than one variable, that may not be known to the researchers. Therefore we conducted an MDS to shed more light on whether there was trust formation.

Appendix E. Multi-Dimensional Scaling (MDS) Using ALSCAL Procedure

There are two ways of conducting the MDS procedure: ALSCAL and PROXSCAL [53]. The ALSCAL procedure involves recording individual scores given by the respondents for every object in the set. Based on these scores, the program computes the distances between objects [52]. We chose the ALSCAL procedure of recording individual scores given by the respondents because the data was individual scores based on the 18 item PTS. There were no distances in our data. Therefore, the variables that we were comparing between the pre-game and post-game results were the differences in the individual PTS scores. The PROXSCAL procedure of collecting the data is recording the distances between objects,

as the respondents perceive them. In this case, the algorithm presents the respondents with all the possible pairs of comparisons [52].

The ALSCAL procedure uses the dissimilarity or discrepancy matrix. The individual scores were aggregated by summation or by computing their average, and:

- If the higher scores correspond to similar objects, and the lower scores correspond to dissimilar objects, then we get a similarity or proximity matrix.
- If on the contrary, the higher scores correspond to dissimilar objects, and the lower scores correspond to similar objects, then we get a dissimilarity or discrepancy matrix [52].

The ALSCAL performed resulted in a dissimilarity matrix, and the chosen model was Euclidean distances with the minimum dimensionality of two, and the maximum dimensionality of two. The SPSS program computed the Euclidean distances between the scores. We chose a two-dimensional model to get one model with two dimensions for ease in the visualization of the conceptual map and interpretation. The maximum model iterations were 30.

The first pieces of information we analyzed were the: (1) model stress; and the (2) stress and squared correlation (RSQ) in distances. They are both indicators of model efficiency. The model stress also called the 'phi' statistics, is the most important. The lower the stress, the better is the model. The critical values to assess model quality are:

1. Stress (phi) *lower than 0.10*, it means that the model quality is *excellent*;
2. Stress (phi) *between 0.10 and 0.20*, it means that the model quality is *good*; and
3. Stress (phi) *greater than 0.20*, it means that the model quality is *poor*.

SPSS computed two values, one proposed by Young (Young's S-stress formula 1) and the other proposed by Kruskal (Kruskal's stress formula 1). Young's S-stress value for the model is 0.0495, and it stopped at the 4th iteration because the S-stress improvement was less than 0.001. Because the value is lower than 0.10, our model is excellent. As for Kruskal's stress, the values were 0.068. Therefore, since both stress values are very low and below 0.10, our model is of excellent quality.

After that, we assessed the RSQ value. RSQ value is "proportion of variance of the scaled data (disparities) in the partition (row, matrix, or entire data) which is accounted for by their corresponding distances" [62]. The RSQ value for the model is 0.98748. Since the RSQ is above 0.95 and close to 1, it confirms that the model is of excellent quality.

Table A5. Clustering of the 36 trust and trustworthy observations based on the derived stimulus configuration from the Euclidean distance model into four clusters, under two dimensions.

Obs. Label	Question/Stimulus Coordinates	Pre-Game Component		Post-Game Component	
		1	**2**	**1**	**2**
TW1	1. Listen(ed) to my conscience	1.12	−0.34	0.89	0.09
TW2	2. Anticipate(d) the needs of others	0.71	−0.06	0.75	0.00
TW3	3. Respect(ed) others	1.17	−0.32	1.29	0.16
T1	4. Gets (got) along with most people	0.59	−0.56	1.15	0.26
TW4	5. Have always been (Was) completely fair to others	0.65	−0.22	0.89	0.20
TW5	6. Believe that laws (game rules) should be strictly enforced	0.88	0.00	0.91	0.09
T2	7. Have (had) a good word for everyone	0.56	−0.25	1.09	0.17
T3	8. Value(d) cooperation over competition	1.10	−0.32	1.17	0.22
TW6	9. Would never cheat on my taxes (never cheated)	0.45	−0.81	1.08	0.05
TW7	10. Follow(ed) through with my plans	0.62	−0.18	1.03	0.20
T4	11. Believe(d) that people (players) are (were) basically moral	−0.07	−0.04	1.02	0.22
TW8	12. Finish(ed) what I start (ed)	0.93	−0.18	1.23	0.09
T5	13. Filled with doubts about things (was filled with doubt)	−1.26	0.21	−2.36	0.34
T6	14. Feel short-changed in life (Felt short-changed)	−1.69	0.47	−2.18	−0.21
T7	15. Avoid contact with other(s) (players)	−1.59	1.11	−2.62	−0.50
T8	16. Believe that most people (players) would lie to get ahead	−0.90	0.78	−2.52	−0.59
T9	17. Find it hard to forgive others (players)	−1.18	1.24	−2.69	−0.38
T10	18. Believe that people (other players) seldom tell the whole story	−0.27	−0.02	−1.95	−0.92
	N: 36				

References

1. Grey, D.; Sadoff, C.W. Sink or swim? Water security for growth and development. *Water Policy* **2007**, *9*, 545–571. [CrossRef]
2. Gleick, P.H. Water and conflict: Fresh water resources and international security. *Int. Secur.* **1993**, *18*, 79–112. [CrossRef]
3. Bakker, K. Water security: Research challenges and opportunities. *Science* **2012**, *337*, 914–915. [CrossRef]
4. Barchiesi, S.; Sanchez, J.C.; Cross, K.; de Madrid, M.P.; Onencan, A.M. Adaptation planning—Views towards resilience and up-scaling success to enhance transboundary water governance (chapter five). In *Transboundary Water Governance: Adaptation to Climate Change*; IUCN Environmental Law Centre: Gland, Switzerland, 2014; Volume 75, pp. 113–241.
5. Enserink, B.; Onencan, A. Nile basin scenario construction. In *IAIA's Contribution in Addressing Climate Change, Proceedings of the 37th Annual Conference of the International Association for Impact Assessment, Montréal, QC, Canada, 4–7 April 2017*; International Association for Impact Assessment: Montréal, QC, Canada, 2017.
6. Onencan, A.M.; Enserink, B.; Chelang'a, J.; Chirchir, W.; Kulei, F. Weshareit Game: Strategic Foresight for Climate-Change Induced Disaster Risk Reduction. In Proceedings of the Humanitarian Technology: Science, Systems and Global Impact 2016, HumTech2016, Boston, MA, USA, 7–9 June 2016; Procedia Engineering: Boston, MA, USA, 2016.
7. Onencan, A.M.; Kortmann, R.; Kulei, F.; Enserin, B. Mafuriko: Design of nzoia basin location based flood game. *Procedia Eng.* **2016**, *159*, 133–140. [CrossRef]
8. Onencan, A.M. Assessment of Hybrid Board Game-Based Learning Outcomes Using the Beatty Theoretical Framework. In Proceedings of the International Simulation and Gaming Association Conference, Delft, The Netherlands, 10–14 July 2017; Springer: Delft, The Netherlands, 2017; pp. 161–172.
9. Onencan, A.M.; Van de Walle, B. Designing disaster diplomacy in the context of a climate change water game. In Proceedings of the International Simulation and Gaming Association Conference, Delft, The Netherlands, 10–14 July 2017; Springer: Delft, The Netherlands, 2017; pp. 43–57.
10. Onencan, A.M.; Van de Walle, B. From paris agreement to action: Enhancing climate change familiarity and situation awareness. *Sustainability* **2018**, *10*, 1929. [CrossRef]
11. Onencan, A.M.; Enserink, B.; van de Walle, B. *Game Design Concept Report: Application of the We Share It Game Elements in Nzoia River Basin*; Delft University of Technology: Delft, The Netherlands, 2018; p. 53.
12. Onencan, A.M.; Van de Walle, B. Equitable and reasonable utilization: Reconstructing the nile basin water allocation dialogue. *Water* **2018**, *10*, 707. [CrossRef]
13. Gambetta, D. Can we trust trust? In *Trust: Making and Breaking Cooperative Relations*; Gambetta, D., Ed.; Basil Blackwell Ltd.: Cambridge, MA, USA, 1988; pp. 213–238.
14. Onencan, A.; Enserink, B.; Wairugala, H.; Chelang'a, J.; Chirchir, W.; Kulei, F. Coupling nile basin 2050 scenarios with the IPCC 2100 projections for climate-induced risk reduction. In Proceedings of the Humanitarian Technology: Science, Systems and Global Impact 2016, HumTech2016, Boston, MA, USA, 7–9 June 2016; pp. 357–365.
15. Strohschein, P. Exploring the Influence of Sand Storage Dams on Hydrology and Water Use. Master's Thesis, University Department in Delft, Delft, The Netherlands, 2016.
16. Hardin, R. *Trust and Trustworthiness*; Russell Sage Foundation: New York, NY, USA, 2002.
17. Mayer, R.C.; Davis, J.H.; Schoorman, F.D. An integrative model of organizational trust. *Acad. Manag. Rev.* **1995**, *20*, 709–734. [CrossRef]
18. Hardin, R. Trust and society. *Compet. Struct.* **2000**, 17–46.
19. Riegelsberger, J.; Sasse, M.A.; McCarthy, J.D. The mechanics of trust: A framework for research and design. *Int. J. Hum.-Comput. Stud.* **2005**, *62*, 381–422. [CrossRef]
20. Uslaner, E.M. *The Moral Foundations of Trust*; Cambridge University Press: Cambridge, MA, USA, 2002.
21. Pich, M.T.; Loch, C.H.; Meyer, A.D. On uncertainty, ambiguity, and complexity in project management. *Manag. Sci.* **2002**, *48*, 1008–1023. [CrossRef]
22. Onencan, A.M.; Meesters, K.; Van de Walle, B. Methodology for participatory gis risk mapping and citizen science for solotvyno salt mines. *Remote Sens.* **2018**, *10*, 1828. [CrossRef]
23. Klinke, A.; Renn, O. A new approach to risk evaluation and management: Risk-based, precaution-based, and discourse-based strategies 1. *Risk Anal. Int. J.* **2002**, *22*, 1071–1094. [CrossRef]

24. Shackley, S.; Young, P.; Parkinson, S.; Wynne, B. Uncertainty, complexity and concepts of good science in climate change modelling: Are gcms the best tools? *Clim. Chang.* **1998**, *38*, 159–205. [CrossRef]
25. Luhmann, N. *Trust and Power*; John Wiley & Sons: Hoboken, NJ, USA, 2017.
26. Ben-Ner, A.; Halldorsson, F. Trusting and trustworthiness: What are they, how to measure them, and what affects them. *J. Econ. Psychol.* **2010**, *31*, 64–79. [CrossRef]
27. Evans, A.M.; Revelle, W. Survey and behavioral measurements of interpersonal trust. *J. Res. Personal.* **2008**, *42*, 1585–1593. [CrossRef]
28. Lee, S.-J.; Ahn, C.; Song, K.M.; Ahn, H. Trust and distrust in e-commerce. *Sustainability* **2018**, *10*, 1015. [CrossRef]
29. Otnes, C.; Lowrey, T.M.; Shrum, L. Toward an understanding of consumer ambivalence. *J. Consum. Res.* **1997**, *24*, 80–93. [CrossRef]
30. Priester, J.R.; Petty, R.E. The gradual threshold model of ambivalence: Relating the positive and negative bases of attitudes to subjective ambivalence. *J. Personal. Soc. Psychol.* **1996**, *71*, 431. [CrossRef]
31. Lewicki, R.J.; Tomlinson, E.C.; Gillespie, N. Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions. *J. Manag.* **2006**, *32*, 991–1022. [CrossRef]
32. Lumineau, F.; Eckerd, S.; Handley, S. Inter-organizational conflicts: Research overview, challenges, and opportunities. *J. Strateg. Contract. Negot.* **2015**, *1*, 42–64. [CrossRef]
33. Rotter, J.B. Interpersonal trust, trustworthiness, and gullibility. *Am. Psychol.* **1980**, *35*, 1. [CrossRef]
34. Reiersen, J. Drivers of trust and trustworthiness. *Int. J. Soc. Econ.* **2018**. [CrossRef]
35. Fukuyama, F. *Trust: The Social Virtues and the Creation of Prosperity*; Free Press Paperbacks: New York, NY, USA, 1995.
36. Rousseau, D.M.; Sitkin, S.B.; Burt, R.S.; Camerer, C. Not so different after all: A cross-discipline view of trust. *Acad. Manag. Rev.* **1998**, *23*, 393–404. [CrossRef]
37. Algan, Y.; Cahuc, P. Trust and growth. *Annu. Rev. Econ.* **2013**, *5*, 521–549. [CrossRef]
38. Gambetta, D. *Trust: Making and Breaking Cooperative Relations*; Blackwell: Oxford, UK, 1988.
39. Pillutla, M.M.; Malhotra, D.; Murnighan, J.K. Attributions of trust and the calculus of reciprocity. *J. Exp. Soc. Psychol.* **2003**, *39*, 448–455. [CrossRef]
40. Bauer, P.C. *Conceptualizing and Measuring Trust and Trustworthiness*; European University Institute: Florence, Italy, 2015.
41. Shubik, M. *Games for Society, Business, and War*; Elsevier: New York, NY, USA, 1975.
42. Geurts, J.L.A.; Duke, R.D.; Vermeulen, P.A.M. Policy gaming for strategy and change. *Long Range Plan.* **2007**, *40*, 535–558. [CrossRef]
43. Zhou, Q.; Bekebrede, G.; Mayer, I.; Warmerdam, J.; Knepflé, M. The climate game: Connecting water management and spatial planning through simulation gaming? In *Water Governance as Connective Capacity*; Ashgate Publishing Ltd.: Farnham, UK, 2013; pp. 109–127.
44. Mayer, I.S. The gaming of policy and the politics of gaming: A review. *Simul. Gaming* **2009**, *40*, 825–862. [CrossRef]
45. Lewis, J.D.; Weigert, A. Trust as a social reality. *Soc. Forces* **1985**, *63*, 967–985. [CrossRef]
46. Lewicki, R.J.; McAllister, D.J.; Bies, R.J. Trust and distrust: New relationships and realities. *Acad. Manag. Rev.* **1998**, *23*, 438–458. [CrossRef]
47. Klein, G.A.; Orasanu, J.E.; Calderwood, R.E.; Zsambok, C.E. *Decision Making in Action: Models and Methods*; This Book Is an Outcome of a Workshop Held in Dayton, OH, USA, 25–27 September 1989; Ablex Publishing: New York, NY, USA, 1993.
48. Klein, G.A. *A Recognition-Primed Decision (rpd) Model of Rapid Decision Making*; Ablex Publishing Corporation: New York, NY, USA, 1993.
49. Garris, R.; Ahlers, R.; Driskell, J.E. Games, motivation, and learning: A research and practice model. *Simul. Gaming* **2002**, *33*, 441–467. [CrossRef]
50. Landers, R.N. Developing a theory of gamified learning: Linking serious games and gamification of learning. *Simul. Gaming* **2014**, *45*, 752–768. [CrossRef]
51. Rotter, J.B. A new scale for the measurement of interpersonal trust. *J. Personal.* **1967**, *35*, 651–665. [CrossRef] [PubMed]
52. Cox, T.F.; Cox, M.A. *Multidimensional Scaling*; CRC Press: Boca Raton, FL, USA, 2000.
53. Kruskal, J.B.; Wish, M. *Multidimensional Scaling*; Sage: Newcastle upon Tyne, UK, 1978; Volume 11.

54. Bachmann, R. Trust, power and control in trans-organizational relations. *Organ. Stud.* **2001**, *22*, 337–365. [CrossRef]

55. Tomlinson, E.C.; Lewicki, R.J. Managing distrust in intractable conflicts. *Confl. Resolut. Q.* **2006**, *24*, 219–228. [CrossRef]

56. Luhmann, N. Familiarity, confidence, trust: Problems and alternatives. *Trust Mak. Break. Coop. Relat.* **2000**, *6*, 94–107.

57. Charnley, G. *Democratic Science: Enhancing the Role of Science in Stakeholder-Based Risk Management Decision-Making*; HealthRisk Strategies: Washington, DC, USA, 2000.

58. Hardin, R. *Trust*; Polity: Cambridge, UK, 2006.

59. Fehr, E. On the economics and biology of trust. *J. Eur. Econ. Assoc.* **2009**, *7*, 235–266. [CrossRef]

60. Reiersen, J. Trust as belief or behavior? *Rev. Soc. Econ.* **2017**, *75*, 434–453. [CrossRef]

61. International Personality Item Pool. A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences. ORI Community and Evaluation Services: Eugene, OR, USA. Available online: http://ipip.ori.org/ (accessed on 7 December 2018).

62. Example of Mds Analyses: Accounting 101 Introduction; International School of Choueifat: Choueifat, Lebanon. Available online: https://psych.unl.edu/psycrs/statpage/mds_eg1.pdf (accessed on 7 December 2018).