

Article

From Paris Agreement to Action: Enhancing Climate Change Familiarity and Situation Awareness

Abby Muricho Onencan *  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.A.vandeWalle@tudelft.nl

* Correspondence: a.m.onencan@tudelft.nl; Tel.: +31-15-27-81-810

Received: 1 May 2018; Accepted: 3 June 2018; Published: 8 June 2018



Abstract: The Paris Agreement was a monumental stride towards global climate change governance. It unlocked the climate change gridlock, introducing country-subjective commitments and a five-year review mechanism. To support the implementation of the Paris Agreement, we designed the Nzoia WeShareIt climate change game. Game sessions were conducted in June and July 2015, and 35 respondents completed a pre- and post-game situation awareness (SA) questionnaire and an in-game performance measurement system. The questionnaire uses a 10-dimensional situation awareness rating technique (SART). Subsequently, we conducted a factorial MANOVA (multivariate analysis of variance) to assess the interaction effects between familiarity, team, and gender. Results indicate an increase in situation awareness. However, policymakers' action was not contingent on the increased SA only, there was a significant interaction effect between familiarity and SA, to lead to climate change actions. Therefore, we recommend more emphasis on the role of familiarity in enhancing SA and, subsequently, supporting the implementation to the Paris five-year review country commitments. We also recommend the increased usage of symbols and capacity development of policymakers on connective capacity to enable them to span the climate change boundaries.

Keywords: climate change disaster risk reduction; symbols; familiarity; policy games; situation awareness; water governance; gender; team cooperation; stories; synecdoche; metaphors

1. Introduction

The Paris Climate Change Agreement that was adopted 12 December 2015 by 195 countries marks a “monumental triumph” [1] that unlocked the “global warming gridlock” [2]. The 1992 UN Framework Convention on Climate Change (UNFCCC) was agreed upon by UN member states so as to reduce greenhouse gas (GHG) emissions [3]. Thereafter, the UN member states sought to develop legally binding rules through the 1997 Kyoto Protocol and subsequent supporting instruments like the Clean Development Mechanism [4]. Unfortunately, despite many policy documents aimed at combating GHG, there was a steady rise in GHG emissions [5]. In 2009, the Copenhagen Conference (COP-15), sought to combat the GHG emissions problem, by creating a subsequent treaty to the Kyoto protocol [6]. The Copenhagen Conference's attempt to police nation states and impose mandatory emission reductions failed [6,7]. At this stage, many observers concluded that the climate change negotiations had reached a deadlock [6]. In the midst of the deadlock, the Paris Agreement was successfully adopted [8]. It is a departure from all previous endeavors to manage GHG emissions. Through the Paris Agreement, countries set their own emission reduction targets. The countries make voluntary pledges and, through regular reviews, poorly performing countries are named and shamed [6]. Therefore, emission cuts are not forced upon countries, but they are voluntarily pledged and later reviewed.

Despite an oversaturated climate change policy regime, the problem of rising GHG emissions still persists [9]. The 2015 Paris Agreement “name and shame” strategy against countries that fail to reduce their GHG emission commitments may not be sufficient in stimulating countries to action. Falkner (2016), states that “time and again, major emitters have shown themselves willing to accept a loss in international reputation when domestic economic priorities have been at stake” [6]. Therefore, there is need to adopt another strategy that may lead countries to take action.

Gupta (2015), assesses the global climate change history and concludes that it has been difficult to create an international climate change regime with mandatory GHG emission targets. The problem is not lack of policies and laws to regulate climate change. According to Falkner (2016), there are numerous global, national and regional climate change policy instruments focusing on institutional capacity development, GHG emission reduction, adaptation, climate-smart agriculture, energy efficiency policies, low-carbon technological innovation and forestry management [6]. Nachmany et al., (2014) states that since 1997, the number of climate change-related laws and policies double every five years. Nachmany et al., (2014) explains that the noted increase in climate change policies applied to both Annex 1 and non-Annex 1 countries. By 2009, there were 426 laws and policies on climate change [10] (p. 20).

Falkner (2016), states that the “past record of climate policies suggests that governments have a tendency to express lofty aspirations but avoid tough decisions” [6]. Giddens (2009), states that the problem lies in the perception of climate change. It is mostly treated as a “back-of-the-mind” [11] problem and future discounted with no present serious action, taken until an actual disaster occurs [12–14]. Falkner (2016) explains the difficulty of translating “normative engagement into collective action” through different facets of the problem [6]. The first facet is an investment problem. Climate change requires extremely high investments in a short duration and the impacts are realized in the long-term. Falkner (2016) adds that it is hard for a government to justify to its citizens high investments with limited or no returns within the short term that the government is in office. The second facet is the complexity of the problem, the climate change effects are not uniform, thus leading to some countries benefiting from the changing climate while others face serious threats. Alexandria, Egypt may face extinction if sea levels rise due to climate change [14–16]. At the same time, Kenya might benefit from the increased rainfall, if they invest more on rain harvesting and water storage facilities [17–20]. The third facet is the deep uncertainty as to whether the climate change projections will materialize [15,18]. Falkner (2016) explains that deep uncertainty has led to governments adopting the “wait and see approach” as the rational decision. Even though climate change is catastrophic, it is seen by many as “abstract and elusive” [11]. According to Luhmann (2000), human beings live within the constraints of a familiar world, familiar risks, familiar hazards, familiar dangers. Luhmann (2000), adds that people never leave the familiar world. Thus, there is a danger of making policies to address only the familiar risks. The fourth facet is the trust issue, most countries lack the certainty that other countries will reciprocate. The diligent work of countries to reduce their GHG emissions might be derailed by “free-riders” who continue business as usual or produce more GHG emissions [6]. Roberts and Parks (2006) explain the last facet, concerning the strained North-South climate change relations. Falkner adds that the tension relates to the division of the mitigation burden. The global south maintains that a large share of the burden should be borne by the industrialized nations who historically are responsible for most of the GHG emissions.

Falkner (2016), explains that there is lack of clarity on whether the Paris Agreement will ensure that climate change commitments translate into action. Before the Paris Agreement, the approach was mainly top-down. For more than two decades countries sought to establish an international climate change regime to impose and enforce mandatory GHG emission reduction targets. However, the Paris Agreement introduced a new climate change governance model. Keohane and Openheimer (2016) state that the Paris Agreement established a “two-level game” approach, linking the global five year Paris Agreement review mechanism to national climate change politics. This “two-level” game is designed to rely on the fear of reputational loss and the availability of nation civil society capacity to

effectively review and advocate for their respective governments to meet their stated commitments. Unfortunately, the “name and shame” strategy against countries that fail to reduce their GHG emission commitments may not be sufficient in stimulating countries to action. Falkner (2016), states that “time and again, major emitters have shown themselves willing to accept a loss in international reputation when domestic economic priorities have been at stake” [6]. Also, many countries lack the civil society capacity and independence to review and advocate for countries to implement their stated commitments.

Paris Commitments may translate into action if the climate change risks are familiar [21]. However, the perception of climate change as a familiar risk that requires immediate action remains a significant challenge [13,17]. Therefore, there is a need to introduce the unfamiliar world, risks, and opportunities into the familiar world and increase policymakers climate change SA.

The goal of this paper is to propose policy recommendations to support the implementation of the Paris commitments and contribute to combating GHG emissions and the climate change governance problem. We propose working with policymakers through introducing the unfamiliar world of climate change to enhance their situation awareness (SA) with the aim of changing their perceptions, attitudes, and behaviors, leading to action. We assess whether increased situation awareness through the introduction of the unfamiliar climate change effects to policymakers may lead to action. We adopt the Endsley (1995) definition of situation awareness as “the perception of the elements in the environment [. . .], the comprehension of their meaning and the projection of their status in the near future” [22]. According to Endsley (1995), there are three levels of SA: level one is perception, level two is comprehension, and the final level is projection. Increased situation awareness depends on the success of level 1 (perception), coupled with comprehension and projection of the change in awareness [22–24]. On the other hand, studies also confirm that the level 1 (perception) of SA is influenced by comprehension and projection [24,25]. Thus, there is a strong linkage between the three levels, and the relationship is not simple nor linear. Bartel (2016) explains that team SA is “the degree to which every team member possesses the situation awareness required for her or his responsibilities” [24] (p. 68).

The study area is Nzoia River Basin in Western, Kenya. The Nzoia river is shared by six county governments, namely Busia, Siaya, Bungoma, Kakamega, Trans Nzoia and Uasin Gishu [26]. The basin area residents are highly susceptible to climate change risks that have in the past led to climate change disasters that include droughts, floods and the activation of the inactive volcanoes along the Great Rift Valley [13,19].

The research results demonstrate that the future may be positive if water policymakers’ awareness of the fragile water system situation is increased, leading to transitions from shortsighted planning towards more long-term and integrated planning that may enable future societies to sustain themselves amidst a changing climate. The results also indicate that familiarity and connective capacity through boundary spanners plays a crucial role in increasing policymakers’ situation awareness, with the aim of changing their perceptions, attitudes, and behaviors, leading to action.

The paper is organized as follows. The theoretical background and analytical framework that supports the research is contained in Section 2. Section 3 describes the research method which comprises of SA assessment, a subsequent factorial MANOVA (multivariate analysis of variance), and the Nzoia WeShareIt game design. The research findings are presented in Section 4 and the discussion in Section 5. The discussion addresses the three core recommendations, triangulation of research methods, the need for capacity development in connective capacity and the policy relevance of bridging the familiarity gap with symbols. Finally, Section 5 makes concluding remarks.

2. Theoretical Background and Analytical Framework

2.1. The Situation: Numerous Climate Change Commitments and Policy Instruments

Kenya actively participates in the climate change international and national arena [27]. On 28 December 2016, Kenya ratified the 2015 Paris Agreement (signed 22 April 2016); which represents a first collaborative virtually unanimous climate change global framework [8]. Prior to that, Kenya signed (12 June 1992) and ratified (30 August 1994), the United Nations Framework Convention on Climate Change (UNFCCC), [3]. Additionally, on 25 February 2005, Kenya ratified the 1997 Kyoto Protocol to the UNFCCC [4]. To implement its international obligations, Kenya developed four main policy documents: the 2016 Climate Change Act that established the National Climate Change Council [28]; the 2010 National Climate Change Response Strategy (NCCRS), [29]; the 2013 to 2017 National Climate Change Action Plan (NCCAP), [30]; and the 2015 to 2030 National Adaptation Plan (NAP) [31]. Kenya has also submitted two reports; the First National Communication (FNC) in 2002 and the Second National Communication (SNC) in 2015 [27] so as to fulfil its obligations under Articles 4 and 12 of the UNFCCC under the guidance of the 1996 Intergovernmental Panel on Climate Change (IPCC) Guidelines [32]. The two national communications (NCs) provide information on greenhouse gas emissions (GHG) in Kenya, including the sources of the emissions and the removals by sinks.

Kenya is not only facing adaptation challenges, increasingly it has to address climate mitigation as well. The second NC indicates that in 2013, Kenya was contributing 0.13 percentage of the GHG emissions global total, which amounts to 60.2 million metric tons of carbon dioxide equivalent (MtCO₂e) [33]. This percentage may seem low when compared to the United States (14.75), the 15 European Union countries (9.33), India (6.43), China (5.93) and Russia (4.86). However, it is a high percentage compared to most developing economies. Kenya is at the same level as Ireland, Serbia, Ecuador and Hungary, whose GHG emissions are also 0.13%. Its GHG emissions are higher than some developed nations, for instance: Denmark (0.12), Switzerland (0.11), Sweden (0.11), Norway (0.10), and Luxembourg (0.02) [34].

Kenya's positive policy reforms have not translated into tangible results with a 59% and 77% increase in agriculture and energy GHG emissions from 1990 to 2013, respectively. As a developing country, its land-use change and forestry (LUCF) sector were expected to be a source of sink to counter the destructive actions of developed countries [35]. Unfortunately, the 1995–2010 GHG inventory indicates that the Kenyan LUCF is a source of emissions [33,36]. The inventory indicates that since 1990 to 2010, Kenyan LUCF activities emitted 17.2 MtCO₂ annually [36]. Agriculture was the most significant contributor of GHG emissions (62.8%), followed by energy (31.2%). The two leading causes of GHG emissions was enteric fermentation from cows, goats, sheep, camels and other ruminant animals that produce methane (55% of the emissions) and manure that has been left on pasture (55% of the emissions). In the energy sector, there were three main GHG emission contributors: transportation (39%), heat and electricity production (26%), and biomass, stationery and mobile combustion (25%) [33,34,36].

To reduce GHG emissions, Kenya has committed to:

- (1) increase the production of solar, geothermal and wind energy;
- (2) improve its energy and resource efficiency;
- (3) increase tree cover to at least 10% of its land mass;
- (4) promote the use of clean energy technologies over wood fuels;
- (5) adopt efficient and low carbon transport modalities;
- (6) the use of climate-smart agriculture (CSA) to implement the Kenya CSA Framework; and
- (7) improve the management of waste [33,36].

Most of these commitments can be initiated through individual and joint action [13,14,17–20,26], without being contingent on financial, investment, technical and human resource support, as stated in the SNC [33,34]. Awareness of the role of government and the community in implementing some of

these actions could improve decision-making with limited or no financial support [14]. For instance, increased awareness of the contribution of the following practices to GHG emissions may lead to positive actions: non-electric vehicles, hydroelectric power production, meat consumption from ruminants, poor land and animal husbandry practices, poor waste management practices, cutting of trees and the increased use of wood fuel [13,17]. This increased awareness may also contribute to informed decision making [18].

Two primary policy instruments have been developed to address the largest GHG emitters: agriculture and energy. To address the GHG emissions from agriculture, Kenya has developed a Climate Smart Agriculture Strategy for the period 2017 to 2026 [37]. The strategy aims at enhancing the adaptive capacity and resilience of pastoralists, farmers, and fisher-folk, reduce GHG emissions, improve institutional collaboration and address CSA cross-cutting issues. In the strategy, some critical issues identified were gender mainstreaming, increased collaboration and improved data and information on CSA [37] (p. 37). To cut down on the GHG emissions from vehicles, an electric Bus Rapid Transit Plus (eBRP+) System for the Nairobi Metropolitan Region is being developed for Nairobi with the support of the World Bank. The first eBRT is known as Ndovu (elephant) BRT line. After completion of the single Greater Nairobi route, the system will be replicated in four following routes: Nyati (buffalo), Chui (leopard), Simba (lion), and Kifaru (rhino), between 2020 and 2030 [38]. These two policy instruments are largely top-down; government-initiated; government-led; and donor dependent.

A policy environment that facilitates the bottom-up adoption of climate-smart agriculture, energy and environmental decisions, is of paramount importance [14,17,20]. Unfortunately, Government-led approaches that are loan or donor financed are not sufficient to address the current climate change challenge. The need for all stakeholders from the public and private sector to embrace individual and group responsibility to address climate change, requires more than top-down, government-led, externally financed, climate change processes and programs.

2.2. The Complication: Future Discounting of Actions Contingent on Climate Change Finance

In accordance with the 2015 Paris Agreement, Kenya's Intended Nationally Determined Contribution (INDC) is the reduction of the GHG emissions by 30 percentage (143 MtCO₂e), relative to the projected business as usual levels by 2030 [33,36]. This commitment is subject to a pre-condition. Kenya made this commitment contingent to receiving financial, investment, technical and human resource support [33,34]. Kenya's 30% reduction in GHG emissions will only translate into action if it received external climate change support to realize the commitment.

Slow or lack of action to address climate change, in Kenya, may lead to water resources management disasters that will affect the quantity and frequency of rain and lead to floods and droughts. According to climate change predictions, Kenya will experience a mean annual temperature increase of between 0.8 °C and 1.5 °C by 2030 and a further increase of between 1.6 °C to 2.7 °C by 2060. In addition, there will be an increment in the frequency by 19–45% and 45–75% of hot days and nights, respectively. The increase in hot days and nights will lead to a subsequent decrease in "cold" days and nights [27]. The Government of Kenya (2015), Second National Communication to the United Nations Framework Convention on Climate Change states that: "cold days and nights are expected to become very rare" [27] (p. 4). The government of Kenya (2015), further adds that there will be a change in the amount of rainfall. The percentage of rainfall change is unknown due to disagreements between various climate change models. The projected rainfall change will range between a decrease of 5% to an increase of 17% by 2030. The most significant increase will be in the months of October to December, and by 2060 there will be a 26% increment in rainfall. This information is useful in the management of water resources. With the projected increase in rain and droughts, there is need to trap and store the water during the rainy seasons and conserve the water for the prolonged drought seasons. This would call for more investments in rainfall trapping and water storage systems. Most of

these actions have been put on hold because the climate change outcomes are uncertain or the planned actions are contingent on climate change financing.

The recent crack along the Great Rift valley in Suswa, Kenya illustrates the assertion that some climate change mitigation actions are on hold until a disaster occurs (Figure 1). On 19 March 2018, Kenya was reported to have split at Suswa along the Great Rift Valley (Figure 1). Kahongeh and Mwangi (2018), attribute the sudden initial split of four Horn of Africa countries from the rest of Africa (Somalia, Kenya, Tanzania, and Ethiopia) to the increased rainfall that washed the volcanic ash, exposing the ash and activating the inactive volcanic activities. Houses were split into half [39], and families vacated their homes before the crack became catastrophic [40]. The tear was more than 15 m deep and 15 m wide [39]. The crack at Suswa had already been projected in previous studies. Skilling (1993) reported the incremental collapse of the Suswa volcano [41]. Bigg et al. (2009), reported multiple inflation and deflation events in a number of Kenyan volcanoes. Suswa was identified to contain active magmatic systems [42] (p. 981).



Figure 1. Aerial view of the extending Great Rift Valley cracks at Suswa, Kenya. Source: Akwei (2018) [39].

The Government of Kenya (2015) report, affirms that the increase during the short rains (March and April) will primarily affect the western Rift Valley region leading to flooding and other climate change-induced disasters [27,31]. The Suswa Rift Valley crack was instigated by a significant flooding event during the March/April rain season.

Risk perception is a crucial component that is required to translate climate change commitments into action [43–46]. Despite numerous studies on the adverse effects of climate change and the effects of heavy rains on the Suswa volcanoes, there has been inaction in the public and private spheres [41,42]. The actual life-threatening tear of the Earth's surface along the Rift Valley, and the subsequent destruction of roads, houses and other infrastructure, increased stakeholders' perception of the particular risk and led to immediate relocation and government action [39,40]. The complication facing Kenya and many other countries is the lack of familiarity of climate change and SA of the need for individual and joint responsibility to address the risks. This has led to inaction, when so much can and still remains to be done, with or without external support.

2.3. The Goal: From Paris Agreement to Action

The goal of this paper is to propose policy recommendations to support the implementation of the Paris commitments and contribute to combating GHG emissions and the climate change governance problem. Through the research we build the capacity of policymakers by introducing the unfamiliar world of climate change to enhance their situation awareness with the aim of changing their perceptions, attitudes, and behaviors, leading to action.

2.4. The Proposed Solution: Enhance Familiarity and Increase Situation Awareness

To arrive at the research goal, we formulated three questions with the use of the research, learning and intervention conceptual framework developed by Mayer, Veeneman (2002) [47] (p. 33).

Learning: Do the policymakers enhance their situation awareness (SA) of climate change risks?

Research: Can increased situation awareness move policymakers from Paris commitments to action?

Intervention: How can familiarity, gender, and team factors contribute to the change in SA of climate change risks in the Nzoia River Basin?

We designed the policy game decision support mechanism using naturalistic decision settings, due to eight factors that complicate climate change decision-making [48]:

1. Ill-structured problems that contain complex causal effects and links;
2. Uncertainty of climate change and the dynamic environment;
3. Constantly changing, poorly-defined, or competing goals;
4. Action/feedback loops—series of events and strings of climate change actions that are intertwined;
5. Time stress when making decisions during the disaster phase;
6. High stakes—large investments and slow returns coupled with deep future uncertainties of the occurrence of the climate change events;
7. Multiple players at multiple levels of governance and from different sectors;
8. National and local government goals and norms need are taken into consideration before making a climate change decision.

Amongst the naturalistic decision models, we adopted the Klein (1993) Recognition-Primed Decisions (RPD) model as illustrated in Figure 2 [49,50]. This model is suitable when the decision maker is an expert (policymaker) and there is time stress (caused by the climate change induced disasters). Klein (1993) explains that policymakers barely undergo an organized decision-making process, where alternatives are assessed when there is time pressure. Policymakers normally assess the nature of the situation and, based on whether the situation is familiar or not, they discount the decision-making, seek more information or proceed to the three phases of decision-making. The first phase is situation recognition, which we renamed perception of climate change elements in their current context. The second phase is serial option evaluation (comprehension), where policy actions are selected from a cue and assessed to select the most typical response. The final phase involves simulating the actions of the policymaker's mind to assess whether they are satisfactory. RPD has two stages where the policy decision may be discounted. First, at the initial stage, if the situation is not familiar. Second, at the perception level, if the expectancies are violated. The focus of this research is to increase the uptake and progression of decision-making by influencing the first stage of the RPD process, where the decision can be discounted for being unfamiliar [49]. The research seeks to increase familiarity so as to enhance the number and quality of policy actions.

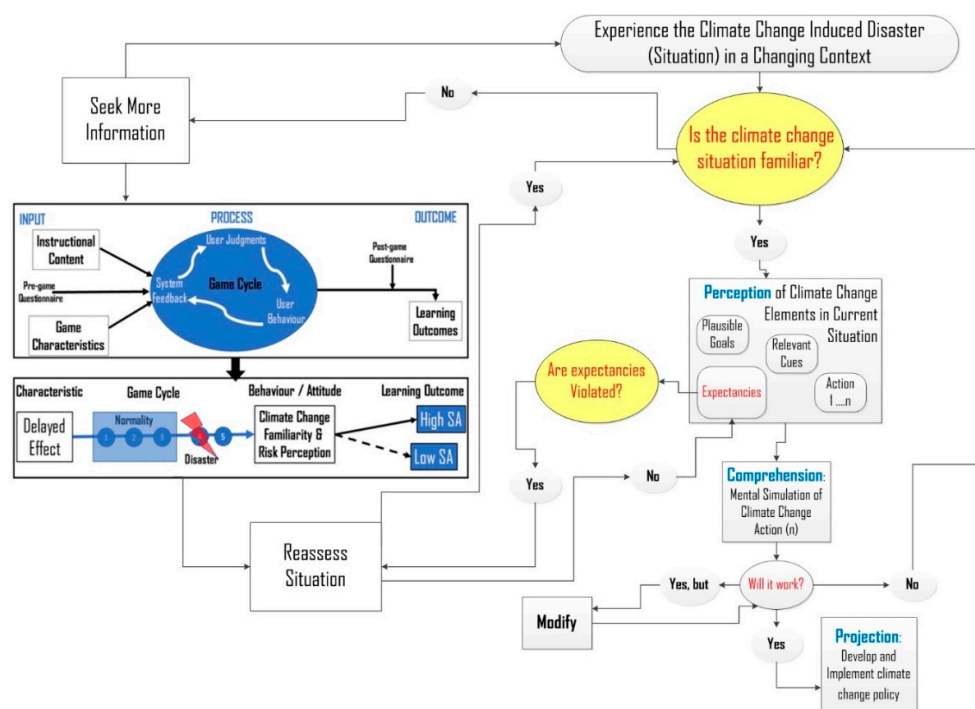


Figure 2. Input-process-output model of Nzoia WeShareIt Game. The arrows in the diagram indicate the gamified learning theory causality path including mediating and moderating processes. Source: Adapted from Klein (1993), [50], Garriss, Ahlers, and Driskell (2002), [51] and Landers (2014), [52].

To be able to design a game that incorporates the SA learning aspect, we combined two design approaches the input-process-output model of serious game design developed by Garriss, Ahlers, and Driskell (2002) and Landers (2014) theory of gamified learning model (Figure 2) [51]. The input-process-output model of serious game design developed by Garriss, Ahlers, and Driskell (2002), was used to design the recurring Nzoia WeShareIt game cycle which comprises steps and cycles. In each cycle, players learn through their actions, judgments (gained while interacting with other players) and the in-game feedback received. The input-process-output model of serious game design was chosen to ensure that the game is designed in a structured way and has been proven in the past to lead to participant learning. In this model, the game instructions and characteristics and the game cycle are vital to the learning process. The instructional content in the game ensured that the players become more familiar with climate change risks and their perception and comprehension of SA are enhanced. The game cycle provides an opportunity for the players to project what they have perceived and comprehended through their strategies and actions. The aim of the input-process-output model of serious game design is to influence the learning process directly.

Landers (2014) theory of the gamified learning model was also incorporated in the design approach because we seek to indirectly influence players' behaviors and attitudes concerning climate change and the perception of unfamiliar risks. Therefore, the enhanced situation awareness rating technique (SART) scores are not sufficient if there is no corresponding change of behavior to support the high SART scores. This model is used to facilitate the process of "digging deeper" into the scores and seek to answer the complex underlying problems of behavior and attitudinal change.

3. Methods

In this section, we provide a summary of the methodological steps used to answer the research questions. There are three steps: first, assess whether the policymakers' situation awareness was increased by comparing the pre-game and post-game SART results [53]. Second, if the SA increases, assess whether it led to action, whether immediate or delayed. Finally, if there was action (whether

delayed or immediate) assess, which of the select three factors (familiarity, gender, and team SA) or their combined effect contributed to the action. We are interested in the interaction effects of the three factors, familiarity, gender and team of players, on situation awareness.

3.1. Situation Awareness (SA)

To assess whether climate change may lead to increased situation awareness, we conducted a quasi-experiment using gaming and simulation. The respondents completed a pre- and post-game situation awareness questionnaire. The questionnaire uses a 10-dimensional subjective pre- and post-trial rating approach developed by Taylor (1990), known as SART. Results indicate an increase in situation awareness on three aspects: (1) demands on attentional resources; (2) supply of attentional resources, and (3) understanding of the situation.

There are several SA measurement techniques, including SART [54]. In the Nzoia WeShareIt game, we used three SA measurement techniques: subjective rating measures (SART) pre-test and post-test questionnaires, performance measures and embedded task measures that were inbuilt in the game [16,26,55]. Endsley (1995) highlights objectivity and less intrusion as critical advantages of performance measures over self-rating subjective techniques. The players did not realize that they were being assessed because the performance matrices were inbuilt in the digital game. Performance measures model a more realistic environment. In addition, performance measures helped in checking the reliability of the subjective SART scores [56]. Another type of performance measurement that was inbuilt in the Nzoia WeShareIt game was external task measures. Sarter and Woods (1991) explain that external task measures entail altering the information and, thereafter, measuring the time taken to react to this change [57]. This measurement technique was introduced in the Nzoia WeShareIt game through a drought round that leads to significant reduction in resources. Endsley (1995) cautions that the technique may be based on wrong assumptions and is highly intrusive. We incorporated this technique because we strategically intended to intrude and alter ongoing tasks and plans thereby disrupting normality with the aim of increasing situation awareness. Team and cross team SA was not included in the design and measurement of the policy game [56]. The Nzoia WeShareIt game also uses the embedded task measures technique to measure situation awareness. The inbuilt electronic game automatically calculated how much trading is done by each participant and what they buy and sell. Also, there is information on how much each player spends on buying food, hydro-electric power, solar power, investing in public services and the payment of penalties. This information can be used to measure many aspects of preferences, strategies, goals and situation awareness levels. The challenge we faced while using the data collected from this technique was the interconnectedness of many factors, that may lead to misleading results. To address this, we used many research techniques to triangulate and confirm the results.

SART is a subjective rating by a person of their level of SA [24]. The technique involves 10 dimensions, based on three 7-point Likert subscales (1 = Low, 7 = High). These subscales measure the degree to which that person perceives (i) the demand on attentional resources (D), (ii) the supply of attentional resources (A), and (iii) the understanding of the situation that they face at that particular moment (U). The factors that comprise demand (D) are the stability of the current situation, the complexity of the situation or and the variability of the situation. Supply of attentional resources (S) includes factors that measure the person's level of concentration and the degree of their spare mental capacity. The factors that influence understanding (U) are the quality and quantity of available information and the extent to which the person is familiar with the situation. According to the SART, SA is measured by combining the ratings in each subscale and then calculating the respondent's SA. Composite SART scores are derived using the following formula:

$$SA = U - (D - S).$$

- U refers to summed understanding.
- D refers to summed demand.
- S refers to summed supply.

In this study, we measure the situation awareness of seven Nzoia policymaking teams. In addition, we contrast this internal perspective with game data derived from the playing of the game by the seven teams, each playing six rounds. The game data collected are the individual scores that each of the SART respondents scored in every round, based on their perception of the game elements in the Nzoia WeShareIt game environment, the comprehension of their meaning in relation to climate change and the projection of their status through long-term planning and joint management of the shared resource. In every game session, we had three facilitators. However, the facilitators did not rate the SA of the policymakers, because of previous research questions about the validity of the rating scores by observers [24,58].

3.2. Factorial Multivariate Analysis of Variance (MANOVA)

We use factorial MANOVA to measure the influence of three independent variables (familiarity, gender and seven teams comprising of five persons each for seven-game sessions) on a dependent variable (situation awareness) with three subscales (demand, supply, and understanding). The researchers selected MANOVA as opposed to analysis of variance (ANOVA) because the tested group differences are on four dependent variables (one SART scale variable for situation awareness and three subscales). Another advantage of using MANOVA instead of ANOVA is its ability to test the differences between the groups on two or more dependent variables simultaneously. It takes into account all the dependent variables and looks at the interaction effect simultaneously. For ANOVA, the analysis is done separately and does not take into account the combined effect.

The three independent variables (also called factors) are categorical, while the dependent variable is continuous. Therefore, the total number of groups compared was 28 ($2 \times 2 \times 7$). The Fisher test was conducted to assess whether the group means for the dependent variable are equal or different.

We designed the three-way ANOVA to study two types of effects: (1) the main effects, this refers to the separate influence of each factor; and (2) the interaction effects, this refers to the combined action of the factors. The study comprises three factors: familiarity, with two (2) levels (low, high); gender, with two (2) levels (male, female); and a team of players, with seven (7) levels (pre and post-game teams for Busia*1, Busia*2, Kakamega*3, Bungoma*4, Bungoma*5, Trans-Nzoia*6 and Trans-Nzoia*7). The research study is designed to assess seven effects in three orders: three main effects: familiarity (F), gender (G) and teams (T) (the separate factor effects); three second-order interaction effects: F*G, F*T, and G*T; and one third-order interaction effect: F*G*T. The detailed design specifications for the MANOVA is contained in Appendix B. Details of the various factorial MANOVA interaction effects studies and the questions and hypotheses that were assessed are contained in Appendix C.

3.3. Nzoia WeShareIt Game Experimental Design

Policy gaming and simulation is one of the most effective approaches used in the recent past to introduce unfamiliar risks. Mayer (2009) defines policy games as:

“Reality is simulated through the interaction of role players using non-formal symbols as well as formal, computerized sub-models where necessary. The technique allows a group of participants to engage in collective action in a safe environment to create and analyze the futures they want to explore. It enables the players to pre-test strategic initiatives in a realistic environment”. [59] (p. 535)

In a game, the risks are low because there are no real-life consequences to the decisions made. As a consequence, games provide policymakers with a safe environment to expose policymakers to possible future climate change risks to steer them to innovative solutions. Therefore, gaming may lead

to policy actions before a disaster occurs as a proactive approach to the challenge, without having to face climate change disasters in a real-life setting [59–62].

To bridge the gap between the familiar and the unfamiliar we need a “s’ymbolon”. “S’ymbolon” is a Greek word for symbols. A symbol is a phrase, object, identity or token, that takes a different meaning or form from the original item or word [63]. Symbols are used to introduce the unfamiliar world into the familiar. The game utilizes climate change disasters to introduce the unfamiliar world and increase the opportunity for planning or taking actions to address unfamiliar risks.

The quasi-experimental design is as follows:

1. The policymakers subjectively rate their situation awareness level before the game using a pre-game questionnaire (low familiarity).
2. During the game, the delayed effect game mechanic introduces a climate change-induced disaster (drought), thus increasing the exposure of the policymakers to risk.
3. The policymakers subjectively rate their situation awareness level after the game using a post-game questionnaire (high familiarity).

3.3.1. The Task

Players in the quasi-experiment were tasked to each manage the resources of a particular county government sustainably and to the satisfaction of their county government residents. The game descriptions for the five county governments are a model of the policy context within the Nzoia River Basin. Players’ strategies and actions require water policy professional abilities used in regular water policymaking. The players receive resources at the start of the game (food, energy, and money) and they were expected to manage these resources in a manner that satisfies their residents’ need for food, energy, and investment in public services.

The game comprises a series of steps and rounds. The steps are the same in all the rounds. However, not all the rounds are the same. The game uses the delayed effect game mechanism to introduce a slow-onset disaster in the form of drought in round four. Based on the player’s strategies, each player survives or fails to survive round 4. If the players do not foresee this climate change risk in the previous rounds and prepare to buffer the system against this risk, then they will not be prepared to avert the devastation that meets them in round 4. In round 4, all their resources are halved. With fewer resources, they are still expected to meet their county government resident needs.

The main problem in round four is food. If a county government is unable to meet the minimum food requirements of their county government residents, then they cannot proceed with the game. Their residents vote them out of office due to their poor short-term policies that failed to take account of deep uncertainties like climate change and buffer the residents from the climate change effects. Also, if the county government faces a significant energy deficit, below the minimum energy need for its residents, the policymaker has to pay the penalty for every energy deficit that they are unable to meet. Finally, in each round, a policymaker is expected to invest in public services. Such investments contribute to the overall score and rating of the policymaker’s performance.

The game does not require policymakers to cooperate to perform their required tasks. However, through different game mechanics (shared goals, complementary information and complementary roles), the game encourages cooperation. The players eventually realize that unilateral actions lead to lower individual results compared to joint actions. One of the game steps is trading. In this step, the players trade their food and energy to be able to increase their income or their food and energy supplies. This game step also encourages cooperation. A detailed description of the game design is provided by Onencan et al. (2018) [26].

3.3.2. Participants and Procedures

Participants in the quasi-experiment are water policymakers from the Nzoia country governments. Onencan et al. (2018) provide a detailed description of the participants and their profiles [26].

There were seven teams; each team played the game on a different day. A team comprised five policymakers, each representing one of the five county governments (Bungoma, Busia, Kakamega, Trans-Nzoia and Uasin Gishu). There were three facilitators from Moi University in Eldoret, Kenya and one gamemaster from the Delft University of Technology.

The quasi-experiments were held in seven separate half-day sessions and conducted at four different Nzoia River Basin county governments (Bungoma, Busia, Kakamega, and Trans-Nzoia). Siaya and Uasin Gishu county governments were not included in the assessment. Uasin Gishu government was involved in the prior game design and testing sessions. Prior to each session, the participants filled in the SART pre-game questionnaire. Immediately after that they were introduced to the game and played six rounds for half a day. After the conclusion of the sixth round, players completed the post-game questionnaire which also incorporated the SART 10-dimension questions. Throughout the quasi-experiment, SA feedback was provided to the players in the form of their game performance scores that were updated real-time and made available on the whiteboard screen, at the end of every round. The in-game leaderboard that was projected on the screen and updated real-time, and there were regular updates from other participants through the step-wise interactions.

3.3.3. Treatments and Measures

The quasi-experiment had no control group. Therefore all the teams experienced the same game environment with the same game mechanics and elements. Each team was exposed to the same treatment conditions of familiarity (low versus high) and mixed gender setting (female versus male). Since it was a quasi-experiment, it could not be treated as a typical 2×2 experimental design. The quasi-experimental treatments and measures are illustrated in Table A4 in Appendix D.

The variables we used (familiarity, gender, and SA) and the constructs they measure are presented in Table 1.

Table 1. Variables definition and measurement.

Construct	Variable	Type	Variable Description	Variable Measurement
Familiarity	The extent of risk exposure	Nominal: dichotomous	The extent to which the policymakers are exposed to climate change risks, which was either high or low (normality)	0 = Low 1 = High
Gender	Interaction of risk familiarity and gender	Nominal: dichotomous	The interaction of the different genders to exposure to high or low climate change risks, amongst the 7 teams (game sessions)	1 = Female 2 = Male
Situation awareness (SA)	Demand	Interval: continuous	Demand level on attentional resources	7 Point Likert Scale
	Supply	Interval: continuous	Supply of attentional resources	7 Point Likert Scale
	Understanding	Interval: continuous	Understanding of the Nzoia Basin situation	7 Point Likert Scale
	Situation awareness	Interval: continuous	Situation awareness that has been attained	Understanding— (Demand–Supply)

As discussed in Section 3.1, SART scores were measured by first deriving the summation of demand, supply and understanding scores. Thereafter, SA was calculated by the use of the following SART formula: $U - (D - S)$. U represents understanding, and D represents demand for attentional resources, and S represents the supply of attentional resources. SART is intended to be measured at the end of the experiment. However, the research approach was designed to measure SART before the start of the game (pre-game) and after the end of the game (post-game). The SART questionnaire was electronically inbuilt in the game and connected to SurveyMonkey so that the results were collected using the SurveyMonkey. In total, seven pre-game and seven post-game teams were assessed. The number of respondents was 70 (35 pre-game and 35 post-game). The observers did not complete

the SART questionnaire. The in-game assessment was different for each player, depending on the county government they were representing. The assessment measured performance based on the amount of food, energy, and investments made, based on five different scales unique for each county government. Based on the policymakers' performance, they collect smileys which accumulate in every round. Onenkan et al. (2018) provide a detailed description of the in-game design and assessment framework. The dataset used to conduct the SA, factorial MANOVA and in-game performance measurements are found in the 4TU repository [53].

4. Results

In this section, we present the experimental findings on the SA of the different water policymakers in the quasi-experiment. Section 4.1, presents the general SA results as well as the underlying dimensions of demand, supply, and understanding. In Section 4.2, the overall in-game results on the individual policymaker's performance are visualized and explained. Section 4.3 focuses on the factorial MANOVA results that assess the role of the three factors (gender, team, and familiarity) on SA. The results section focuses on answering three research questions:

1. Learning: Do the policymakers enhance their situation awareness of climate change risks?
2. Research: Can increased situation awareness move policymakers from Paris commitments to action?
3. Intervention: How can familiarity, gender, and team factors contribute to the change in SA in climate change risks in the Nzoia River Basin?

4.1. Climate Risk Situation Awareness of the Nzoia River Basin Policy Makers

The SA findings are based on subjective SA scores of the Nzoia River Basin policymakers before the start of the Nzoia WeShareIt game (pre-game questionnaire) and at the end of the game (post-game questionnaire).

Each team consisted of five policymakers each representing the five select county governments in the Nzoia River Basin (Bungoma, Busia, Kakamega, Trans Nzoia and Uasin Gishu). Since we had 7 teams, each with 5 members and SA was measured pre- and post-game, we had in total 70 measures of SA.

The results indicate an increase in SA at all levels (demand, supply, and understanding), as illustrated in Table 2. The standard deviation scores indicate a spread out of the scores at the pre-game stage and more convergence towards the mean at the post-game stage. Table 2 summarizes the means, standard deviations, and percentiles of the policymakers' situation awareness. Table A3 provides the descriptive statistics for the familiarity factor, in Appendix D.

Table 2. Descriptive statistics of SA and its dimensions (demand, supply, and understanding) for the pre-game and post-game findings (N = 35).

	N	Mean	Std. Deviation	Min	Max	Percentiles		
						25th	50th (Median)	75th
Pre-Game Demand	35	8.23	4.14	3.00	18.00	5.00	8.00	10.00
Post-Game Demand	35	16.69	3.22	10.00	21.00	14.00	17.00	20.00
Pre-Game Supply	35	9.94	5.37	4.00	28.00	6.00	9.00	12.00
Post-Game Supply	35	24.17	3.67	13.00	28.00	22.00	25.00	27.00
Pre-Game Understanding	35	8.43	3.88	3.00	20.00	6.00	8.00	11.00
Post-Game Understanding	35	19.37	2.14	14.00	21.00	18.00	20.00	21.00
Pre-Game SA	35	10.14	7.88	−3.00	38.00	4.00	9.00	13.00
Post-Game SA	35	26.86	5.21	17.00	38.00	23.00	27.00	30.00

We checked for significant outliers. According to the results, there was only one significant outlier as illustrated in the boxplot diagram in Figure 3. The respondent that was identified as an outlier had extremely high pre-game SA scores compared to other players in all the seven teams. Since it was only one outlier, we decided to keep the respondent results in the subsequent analysis.

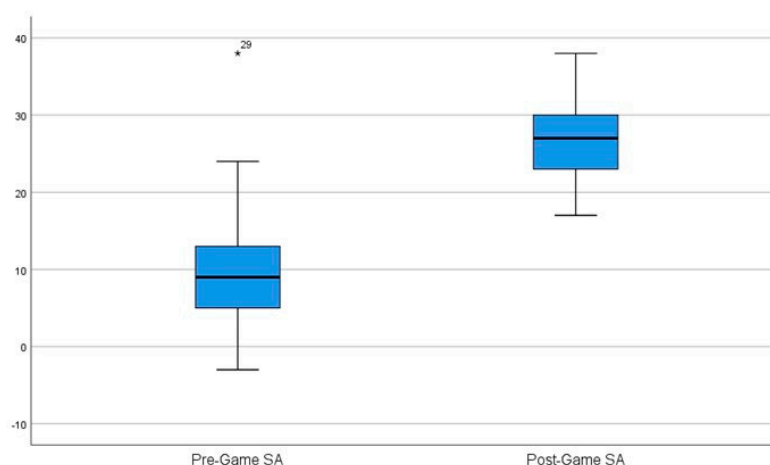


Figure 3. Boxplot to identify significant outliers. The Y-Axis represents the total SA score for each of the 35 respondents whereas the X-Axis visualizes two separate statistics for the two data subsets, the individual SA scores for the pre-game and post-game.

An ANOVA using Friedman's test and Tukey's test for non-additivity for SA scores was conducted. The ANOVA shows that there is a statistically significant increase in situation awareness at the $p < 0.05$ level, $F(1, 34) = 26.85, p = 0.005$. The ANOVA test details are in Table A1 (Appendix A). The increase is also visualized in the Boxplot (Figure 3).

4.2. The Contribution of Nzoia WeShareIt Policy Game to Enhancing SA

The in-game findings indicate a cumulative improvement in game performance with team 1 and 2 being the least performing teams and team 5, 6 and 7 being the best performing teams (Figure 4). Results show that there was cross learning within and between teams. The within-team learning is demonstrated by the improved results after every successive round. The between-team cross-learning is demonstrated by improved overall performance and the mastering of the game after each successive game session.

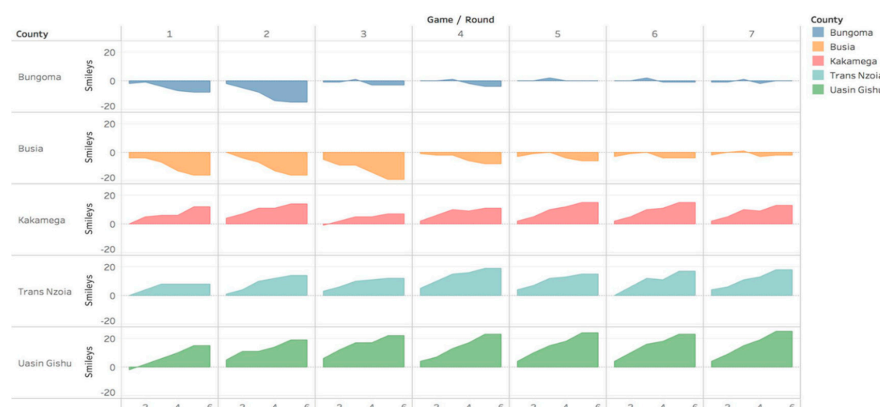


Figure 4. Smileys received by local governments per round—the extent to which they made their citizens happy through the provision of food, energy, a healthy environment and investment in public services.

It is important to note that the players did not meet each other before and after the game sessions. Therefore, we concluded that the boundary spanners that connected the different teams so as to enhance the between-team SA, were the three facilitators. There was no quantitative data that was collected to verify this assumption. The assumption on the contribution of the facilitators as boundary spanners to

the 7 teams is based on the qualitative data captured in the game observations, the rough-cut game video recordings, and the debriefing session notes.

4.3. Factorial MANOVA Results

A 3×4 factorial MANOVA was conducted to compare the effect of three independent variables (IDVs), (gender, familiarity, and team) on the overall situation awareness as well as on the three SA dimensions (demand, supply, and understanding). Table A2 (Appendix B) lists effects, questions, and hypotheses for the three IDVs (gender, team, and familiarity) and their interaction effects. Table A3 contains the main descriptive statistics.

The highest order interaction effect (the third order interaction effect), indicates a significant difference between the levels of familiarity levels (high or low), gender (female or male) and teams (one of the seven teams), differ when considered jointly on the variables demand, supply, understanding and situation awareness, (Wilk's $\Lambda = 2.82$ (F 9, 112.10) = 0.78, $p = 0.01$, partial $\eta^2 = 0.15$).

The results of the MANOVA indicated that there is no significant simple second-order interaction effect (Table A5). In particular, there was no significant difference between policymakers with different familiarity levels (high or low) and gender (female or male), when considered jointly on the variables demand, supply, understanding and situation awareness (Wilk's $\Lambda = 0.95$ (F 3, 46) = 0.82, $p = 0.49$, partial $\eta^2 = 0.05$). The results also indicate no significant difference between policymakers with different familiarity levels (high or low) and teams (one of the seven teams), when considered jointly on the variables demand, supply, understanding and situation awareness (Wilk's $\Lambda = 0.61$ (F 18, 130.59) = 1.37, $p = 0.16$, partial $\eta^2 = 0.15$). In addition, there was no significant difference between policymakers with different gender (female or male) and teams (one of the seven teams), when considered jointly on the variables demand, supply, understanding and situation awareness (Wilk's $\Lambda = 0.94$ (F 9, 112.10) = 0.30, $p = 0.97$, partial $\eta^2 = 0.02$).

For the 1st main effect, the results of the MANOVA indicated that there was a significant difference between high and low familiarity on the three subscales and the overall SA, (Wilk's $\Lambda = 0.16$ (F 3, 46) = 82.74, $p = 0.005$, partial $\eta^2 = 0.84$). For the 2nd main effect, there was no significant difference between female and male SART scores on the three subscales and the overall SA, (Wilk's $\Lambda = 0.95$ (F 3, 46) = 0.78, $p = 0.51$, partial $\eta^2 = 0.05$). For the 3rd main effect, there was no significant difference between high and low familiarity on the three subscales and the overall SA, (Wilk's $\Lambda = 0.75$ (F 18, 130.59) = 0.76, $p = 0.74$, partial $\eta^2 = 0.09$).

The 2×4 MANOVA assessment of the familiarity IDV indicates a significant difference between the group means for the familiarity (F) factor on the overall situation awareness as well as on the three SA dimensions (demand, supply, and understanding) as the dependent variables (Table A5).

Follow up tests of between-subjects effects were conducted for gender, team, and familiarity, with each ANOVA conducted at an alpha level of 0.05. The results confirm the MANOVA results. There was no significant gender or team between-subjects effect on the overall SA and its three dimensions (demand, supply, and understanding). The detailed results of the test of between-subjects effects can be found in Table A6.

The tests of between-subjects effects indicate significant demand effect (F 1, 6) = 92.27, $p = 0.005$, partial $\eta^2 = 0.66$), with the post-game SART results reporting significantly higher familiarity on demand for attentional resources than the pre-game SART results.

The familiarity between-subjects tests also indicated a significant supply effect (F 1, 6) = 148.81, $p = 0.005$, partial $\eta^2 = 0.76$), understanding effect ((F 1, 6) = 180.26, $p = 0.005$, partial $\eta^2 = 0.79$), and overall SA effect (F 1, 6) = 112.96, $p = 0.005$, partial $\eta^2 = 0.70$).

The post-game SART results reported significantly higher familiarity effects on all the four dependent variables (overall SA, demand on attentional resources (D), the supply of attentional resources (A), and the understanding of the situation that they face at that particular moment (U)), than the pre-game SART results.

Follow up univariate tests of between-subjects effects were also conducted for third and second order interaction effects of gender * familiarity * team (third-order effect), familiarity * gender (2nd-order effect), familiarity * team (2nd-order effect), and gender * team (2nd-order effect). The results confirm the MANOVA results. There was no significant second order interaction effect. However, the third order interaction effect indicates mixed results (Table A6).

The mixed results indicate significant combined gender * familiarity * team (third-order effect) interaction effect on demand ($F(3, 68) = 5.57, p = 0.005$, partial $\eta^2 = 0.26$) and the overall SA ($F(3, 68) = 4.56, p = 0.01$, partial $\eta^2 = 0.22$). The third order interaction effect was not significant on supply ($F(3, 68) = 1.53, p = 0.22$, partial $\eta^2 = 0.09$) and understanding ($F(3, 68) = 0.89, p = 0.89$, partial $\eta^2 = 0.01$).

The post-game SART results reported significantly higher familiarity effects on all the four dependent variables (overall SA, demand on attentional resources (D), the supply of attentional resources (A), and the understanding of the situation that they face at that particular moment (U)) than the pre-game SART results. The mean difference between low and high familiarity is 16.71. Table A5 details the results of the tests of between-subjects effects for familiarity factor. Table A6 contains the univariate test results that test the effects of familiarity, based on pairwise comparisons (Table A9) that are linearly independent among the estimated marginal means.

Overall SA has an important influence on the dependent variable familiarity. Additionally, D and SA have a significant influence on the third-order combined effect of F^*G^*T . To assess how big is the influence of SA on the dependent variable familiarity we assessed the difference between the groups by consulting the table of pairwise comparisons (Table A9, see also Tables A8 and A10 in Appendix D). To demonstrate that SA increases familiarity, we maintained only the positive difference. Therefore, the mean difference between the low familiarity and the high familiarity groups is 16.71; the p -value is lower than 0.0005. Thus the difference is statistically significant. In conclusion, SA is effective at high climate change risk familiarity levels.

In summary, we noticed that both familiarity and the third order combined effect of F^*G^*T effect grow when the SA increases. These findings reveal the importance of familiarity in enhancing SA at all levels. While an increase in familiarity leads to a subsequent increase in demand, supply, understanding, and overall SA, the third-order combined effect of F^*G^*T only affected demand and the overall SA. For the gender and team factor to have any effect, they need to be combined with familiarity at the third order effect level. Any lower level (lower than the third level of interaction) interaction where familiarity is not incorporated in the factors, led to no significant results. In addition, it is not clear whether the third-level interaction was only significant because of familiarity because the other two factors do not seem to have any effect at the second-order and main effect levels. Therefore, we conclude that familiarity is a critical factor that should be incorporated into the design and implementation climate change risk situation awareness interventions.

5. Discussion: Key Findings and Policy Relevance

5.1. Summary of the Key Research Findings

The research results can be summarized in three main findings:

1. There was a significant increase in player SA when comparing the pre-test and post-test SART results. The pretest individual scores were treated as the baseline data. The movement from commitments to actions is a complex socio-technical system that requires further analysis. We, therefore, propose triangulation of the research measurement method to effectively assess this complexity (see Section 5.2.).
2. We noticed that increased SA did not lead to immediate actions. Actions were only taken by the later teams after hearing stories on previous game sessions from the facilitators. Therefore, there are two key elements to successful policy implementation: a story (see Section 5.4) and

a person with the connective capacity to effectively narrate the story and span the boundaries between two or more geographically dispersed teams (see Section 5.3).

3. The results indicate that increased SA only leads to action if the policymakers are familiar with climate change actions and there is a combined interaction effect between gender, team (mainly cross-team) and familiarity. To ensure gender balance, we recommend mainstreaming gender in climate change processes and actions. Gender mainstreaming will be addressed in more detail in a subsequent publication. For the team, we recommend more capacity development of policymakers' connective capacity to enable them to span the multiple climate change boundaries. Team-interdependence and social learning will also be addressed in more detail in subsequent publications. For familiarity, we recommend an increase in the quantity and quality of climate change stories, metaphors, and synecdoches as explained in Section 5.4.

5.2. Triangulation of SA Measurement Techniques for Enhanced Policy Game Insights

The findings in Section 4.1, indicate an unexplained variance between the subjective individual SART scores and the game results. The SART results measure individual situation awareness and do not take into account team and cross-team SA. These findings are based on the three SA measurement techniques that we used: subjective rating measures (SART) pre-test and post-test questionnaires, performance measures and embedded task measures that were inbuilt in the game. The findings reveal positive influence of within-team and cross-team SA. There was cross-learning between these geographically dispersed teams and they did not have any contact during the game sessions. Based on the in-game findings, there was a cumulative improvement in game performance with team 1 and 2 being the least-performing teams and team 5, 6 and 7 being the best-performing teams. This indicates that there was a form of social learning that kept building up with each successive round and game, despite the weak linkages between the teams, if any.

In similar future research, game designers should triangulate a number of measurement techniques. Apart from the three measurement techniques used in the research, game designers should also consider using observer-ratings to test other factors that influence the final results and the in-game freeze technique. The observer-rating technique was not incorporated in the Nzoia WeShareIt assessment. However, it proved useful from the game results. The game outcomes indicate that the facilitators had a significant influence on the player's SA. This conclusion could only be inferred because the facilitators were not incorporated in the measurement techniques. The observer-rating technique requires an independent, knowledgeable observer to rate the SA of the players and the facilitators. This observer could also assess the role of the facilitators in the team and cross-team cooperation when the teams are dispersed. Endsley (1995) explains that the freeze technique involves random freezing of the system displays and suspending simulations for a short moment to allow the participants to reflect on the perception of the situation [56]. This approach is implemented several times during the simulation. We noticed that half a day was too long before reflection and so many things happen and are forgotten during the game session. The debriefing was not useful in measuring situation awareness, especially just before lunch when the participants are hungry or plan to return back to their respective offices. Therefore, the freezing technique would be ideal for addressing some of these challenges.

Triangulation is the proposed approach to ensure more objective, reliable and valid results that can easily be tested and confirmed with a separate set of results measured on the same respondents during the same climate change gaming simulation. Many techniques can be used to measure the enhancement of climate change situation awareness. Each situation is different. Therefore, one technique might work in one case study and not in another. Policy game designers should understand the contribution, value, and drawbacks of each technique, before finally selecting the suitable set of techniques.

5.3. Role of Boundary Spanners in Enhancing Climate Change Governance

The policymakers within the 5 teams' increased their SA by actively participating in the policy game. Unfortunately, the in-game data indicates that increased SA was not sufficient to spur policymakers to undertake policy action. The results indicate the need to hear actual stories from someone who had experienced that game in a previous game session in order to take action. The first players did not have this advantage, and thus they were not able to implement what they had learned quickly.

For the climate change discourse to change people and for these changed individuals to take action, there is a need for boundary spanners. A boundary spanner enters an unfamiliar world, experiences the unfamiliar world, and comes back, to the familiar world with unfamiliar experiences. As such, the facilitators' stories of previous game sessions were symbols that made climate change risks and opportunities not unfamiliar to the new team of policymakers. Through the stories of experiences in the previous game sessions, the policymakers were ready to take the risk of moving from commitment to action aimed at addressing climate change risks. This change happened, because the risks no longer were unfamiliar.

Future climate change interventions should incorporate boundary spanners, to spur change from within the system through horizontal social networks. Climate change boundary spanners can play the following roles:

1. Unfamiliar climate change information processing and validation through experience;
2. The external representation of the dynamic climate change system that they have experienced to persons who are still unfamiliar with the climate change risks and opportunities;
3. Monitoring climate change-related impacts, projects, and opportunities;
4. Scanning the system for climate change risks and opportunities; and
5. Acting as climate change gatekeepers.

5.4. Policy Relevance: Bridging the Familiarity Gap with Stories, Synecdoches, and Metaphors

The research findings indicate that familiarity plays an extremely significant role in instigating policy action to support the implementation of the Paris Commitments. However, there is little guidance on how to introduce the unfamiliar climate change world into the current familiar world. Stone (2002) defines a symbol as "anything that stands for something else" [63] (p. 157). Stone (2002) explains that symbols may seem trivial, but they have the ability to take living form, which is not possible with climate change facts and numbers. Symbols are used to represent an unfamiliar world within the familiar world. Once they take a living form of their own, then the unfamiliar world ceases to be unfamiliar [63].

Policy gaming is a useful tool that can be used to introduce climate change symbols in the form of (1) stories; (2) synecdoche's; and (3) metaphors. Climate change stories are narratives of climate change villains and heroes, risks, and opportunities; problems and solutions; and resolutions and tensions, introduced in a storyline. Climate change synecdoches represent the whole with only a small part. A useful synecdoche are horror stories of climate change-induced disasters. The cracking of Africa into two along the Kenyan Great Rift in Suswa is a horror story and if used well can be a successful strategy to initiate and maintain climate change actions. Climate change metaphors are used to liken one policy problem to the climate change problem. Some of the common metaphors that can be likened to climate change are climate change-related diseases, climate change and water crisis, climate change wars, climate change refugees and immigrants and climate change natural disasters. Metaphors are useful in bridging the climate change familiarity gap. Although climate change may seem unfamiliar, linking it with current policy problems that are considered "real" gives life and form to the climate change story.

Successful climate change story-making requires a careful balance between the two sides of the narrative. First, the story should contain two sides of the narrative. Second, the two sides should be balanced. Game simulation of two sides of the story and ensuring there is a balance between the two

sides are critical competencies that all climate change game designers should have. Stories of power must contain two sides: helplessness and control. If the helplessness aspect is too strong that it clouds the control part, then the story is not balanced and is easily discounted as an illusion. Most of the climate change stories leave the listeners feeling helpless with no sense of control. That is why they are barely taken into account as real stories that necessitate action. Some climate change stories also take the form of “random”, “accidental”, “deeply uncertain”, “natural”, and “a twist of fate” [63] (p. 166). The imbalance in the narrative makes it difficult for the story to lead to action. Story imbalance leaves listeners feeling helpless, leading to minimal or no action.

Policy gaming could play a prominent role in the development and narration of balanced climate change storylines, metaphors, and synecdoches. A critical aspect of the Nzoia WeShareIt game was the introduction of climate change synecdoche’s in the form of drought. However, this drought horror must be balanced with the positive opportunities that arise out of the disaster to avoid leaving the players feeling helpless.

Human agency and how individual, societal and state actions can bring about positive change is a critical element that should always be considered when crafting a storyline. However, caution should be taken not to tip the balance towards human agency. Climate change stories that are heavily skewed towards human agency, take the form of conspiracy theories. These stories create the impression that climate change reforms can only be done by a few influential people. Conspiracy stories leave the listeners powerless, and no action is taken. Alternatively, there are also stories that confine human agency to a select few. One prominent story that has taken center-stage in the climate change negotiations for more than two decades is the blame-the-victim story. The Western worlds are blamed for destroying the ozone layer, and thus they should pay reparations to developing countries. Such stories blind the developing countries to the many actions happening within their boundaries that may be contributing to the global climate change crisis. As indicated in this research, Kenya is significantly increasing its GHG emissions but still maintains that its commitments are contingent on external climate change support (human, financial, investment and technological). This story precludes the Kenyan government and citizens who continue to buy GHG-emitting vehicles, destroy forests or keep large herds of cattle that emit methane, from taking immediate steps to reduce the current GHG emissions. Blame the victim stories create an “us versus them” mentality where the Western world denies culpability and the developing world waits upon the Western world to fix the problem. Most important to the climate change discourse are the stories of change. Stories of change consist of two sides: decline and progress. A careful balance between the current climate change decline story and the progress made to address and curb the decline is of critical importance. Future research should focus on progress stories, to tilt the scale away from the decline stories, towards a more balanced narrative.

Future climate change research should assess the contribution of symbols in the climate change policy discourse, specifically in the following research fields:

1. Contribution of climate change stories of progress on increased SA;
2. Creation of alliances around a climate change policy problem with the aim of developing shared meaning;
3. Reducing stories that promote helplessness and supporting societies to gain control and strengthen climate change bottom-up, goal-oriented movements and interest groups;
4. Enhancing climate change policymaking that facilitates bottom-up implementation;
5. Encouraging climate change collective action at the local, national, regional and global level [63] (p. 181).

6. Conclusions

The adoption of the Paris Agreement was a great stride in global climate change governance. The agreement changes the traditional mode of international cooperation that is mainly top-down to a hybrid model that incorporate both top-down and bottom-up elements. The core mechanism of the

Paris Agreement is the five-year review mechanism. However, studies indicate that this mechanism is bound to fail, if not supported.

The paper proposes policy recommendations to support the implementation of the Paris commitments and contribute to combating GHG emissions and the climate change governance problem. Through research, we build the capacity of policymakers by introducing the unfamiliar world of climate change to enhance their situation awareness with the aim of changing their perceptions, attitudes, and behaviors, leading to action. Three research methods were used. SART for the pre-test and post-test subjective ratings of SA. In-game performance measures assessed whether the increased SA led to the implementation of policy actions. Finally, MANOVA assessed the interaction effects of familiarity, gender, and the 5-member teams. To arrive at the research goal, we formulated three questions:

Learning: Do the policymakers enhance their situation awareness (SA) of climate change risks?
 Research: Can increased situation awareness move policymakers from Paris commitments to action?
 Intervention: How can familiarity, gender, and team factors contribute to the change in SA of climate change risks in the Nzoia River Basin?

The research results can be summarized in three main findings:

1. There was a significant increase in player SA, between the pre-test and post-test SART results. However, the movement from commitments to actions is a complex socio-technical system that requires further analysis through the use of triangulation.
2. There are two key elements to successful policy implementation: a story and a person with the connective capacity to effectively narrate the story and span the boundaries between two or more geographically dispersed teams.
3. Increased SA only leads to action if the policymakers are familiar with climate change actions and there is a combined interaction effect between gender, team (mainly cross-team) and familiarity.

The overall research findings indicate that familiarity plays a significant role in instigating policy action to support the implementation of the Paris Commitments. However, there is little guidance on how to introduce the unfamiliar climate change world into the current familiar world. We recommend that there should be more focus on the role of symbols in facilitating the change towards implementing the Paris Commitments. For the stories, we recommend that they should be balanced, have two sides and ensure that human agency is promoted.

Future climate change research should assess the contribution of symbols in the climate change policy discourse. In summary, the climate change discourse can be changed through the use of boundary spanners and symbols to bridge the familiarity divide between what is considered real or merely an illusion.

Supplementary Materials: 1. The DOI of the dataset used to conduct the SA, factorial MANOVA, and in-game performance measurements are found in the 4TU repository: 10.4121/uuid:b7f561b1-b2e1-468c-a653-069af41b65e3
 2. The DOI of the detailed game design concept report is found in the TU Delft Repository: <http://doi.org/10.4233/uuid:3a1f1e27-274b-4459-8a14-ba3068a9cb4a>.

Author Contributions: A.M.O. conceptualized the article, customized the SART questionnaire for the Nzoia river basin, designed the Survey Monkey questionnaire, conducted the game sessions, collected the data, undertook the in-game performance measurements analysis and the factorial MANOVA in the context of SA, wrote the original draft and was actively involved in the draft preparation, content visualization, draft improvement and the incorporation of comments from the second author and the reviewers. B.V.d.W. was actively involved in the policy game design and the game-testing sessions, mobilized the financial and technical resources to design and implement the policy game, improved the initial conceptualization, methodology, was actively involved in the validation process, and was also responsible for resources, review & editing, supervision, project administration and formal analysis.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Situation Awareness Analysis of Variance (ANOVA) Results

We conducted an ANOVA using Friedman's test and Tukey's test for non-additivity for SA scores. The ANOVA shows that there is a statistically significant increase in situation awareness at the $p < 0.05$ level, $F(1, 34) = 26.85$, $p = 0.005$. The ANOVA test details are in Table A1 (Appendix A).

Table A1. ANOVA with Friedman's Test and Tukey's Test for Nonadditivity for SA.

ANOVA with Friedman's Test and Tukey's Test for Nonadditivity							
		Sum of Squares	df	Mean Square	Friedman's Chi-Square	Sig.	
Between People		1549.00	34	45.56			
Within People	Between Items	4888.93 ^a	1	4888.93	26.85	0.00	
	Residual	Nonadditivity	227.78 ^b	1	227.78	5.99	0.02
		Balance	1255.79	33	38.05		
		Total	1483.57	34	43.63		
	Total		6372.50	35	182.07		
Total		7921.50	69	114.80			
Grand Mean = 18.5000							

^a Kendall's coefficient of concordance $W = 0.617$. ^b Tukey's estimate of power to which observations must be raised to achieve additivity = 1.849.

Appendix B. Design of the Factorial MANOVA Model

We designed the three-way ANOVA to study two types of effects: (1) the main effects, this refers to the separate influence of each factor; and (2) the interaction effects, this refers to the combined action of the factors.

The study comprises three factors:

1. familiarity, with two (2) levels (low, high);
2. gender, with two (2) levels (male, female);
3. a team of players, with seven (7) levels (pre and post-game teams for Busia*1, Busia*2, Kakamega*3, Bungoma*4, Bungoma*5, Trans-Nzoia*6 and Trans-Nzoia*7).

The research study is designed to assess seven effects in three orders:

1. three main effects: familiarity (F), gender(G) and teams (T) (the separate factor effects);
2. three second-order interaction effects: F*G, F*T, and G*T;
3. one third-order interaction effect: F*G*T.

The simple main effects of any factor represent the influences of that factor when the levels of the other factors remain unchanged. To illustrate this, the number of groups in the simple main effects model for the factor familiarity (low or high) is 28 ($2 \times 2 \times 7$), as listed below:

1. Seven groups measuring the effect of low familiarity on male policymakers for each of the seven teams (Busia*1, Busia*2, Kakamega*3, Bungoma*4, Bungoma*5, Trans-Nzoia*6, and Trans-Nzoia*7);
2. Seven groups measuring the effect of low familiarity on female policymakers for each of the seven teams (Busia*1, Busia*2, Kakamega*3, Bungoma*4, Bungoma*5, Trans-Nzoia*6, and Trans-Nzoia*7);
3. Seven groups measuring the effect of high familiarity on male policymakers for each of the seven teams (Busia*1, Busia*2, Kakamega*3, Bungoma*4, Bungoma*5, Trans-Nzoia*6, and Trans-Nzoia*7); and
4. Seven groups are measuring the effect of high familiarity on female policymakers for each of the seven teams (Busia*1, Busia*2, Kakamega*3, Bungoma*4, Bungoma*5, Trans-Nzoia*6 and Trans-Nzoia*7).

The simple second-order interaction effects represent the interaction effects of two factors for each level of the third factor. The simple third-order interaction effects represent the sum of all the interaction effects between the three factors.

The following assumptions form the basis of the analysis:

1. The three independent variables are categorical, each having at least two categories.
2. The dependent variable is continuous.
3. Observations are independent; there is no relationship between the subjects in our groups.
4. The dependent variable is normally distributed in all groups.
5. The dependent variable does not present significant outliers in any group.
6. The dependent variable has equal variances in all groups (variances are homogeneous).

We tested all the assumptions, and the results were positive except the Box's test of equality of variance. This test seeks to check whether the covariance matrices of the dependent variables are equal across groups. Box's test of equality of covariance matrices was not computed because there are fewer than two nonsingular cell covariance matrices. Since the SART scale is an already established scale and it has been tested and verified in many studies, we decided to proceed with the analysis without any results from the Box's Test of Equality of Covariance.

Appendix C. Factorial MANOVA Effects, Questions, and Hypotheses

Table A2 illustrates: (1) the MANOVA effects by familiarity (0,1) team (1,7) gender (1,2), (2) the questions to be answered and (3) the tested hypotheses. All dependent variables are related to the SART subjective rating technique for demand, supply, understanding and situation awareness.

Table A2. MANOVA effects by familiarity (0,1) team (1,7) gender (1,2), the questions to be answered and the tested hypotheses. All dependent variables related to the SART subjective rating technique.

Effects Tested	Question to be Answered	Hypotheses (the mean difference is significant at the 0.05 level), [N = 70 cases]
1st main effect	Does high and low familiarity differ on the three subscales and the overall scale of SA?	<ol style="list-style-type: none"> 1. H1a: the group means for the familiarity (F) factor are equal in the total population. 2. H1b: the group means for the familiarity (F) factor are different in the total population.
2nd main effect	Do male and female policymakers differ on the three subscales and the overall scale of SA?	<ol style="list-style-type: none"> 3. H2a: the group means for the gender (G) factor are equal in the total population. 4. H2b: the group means for the gender (G) factor are different in the total population.
3rd main effect	Is there a difference between the seven (7) teams on the three subscales and the overall scale of SA?	<ol style="list-style-type: none"> 5. H3a: the group means for the teams (T) factor are equal in the total population. 6. H3b: the group means for the teams (T) factor are different in the total population.
1st second-order effect	Do policymakers with different familiarity levels (high or low), and gender (female or male) differ when considered jointly on the variables demand, supply, understanding and situation awareness?	<ol style="list-style-type: none"> 7. H4a: the sum of all the interaction effects between factors F and G, is equal to zero. 8. H4b: the sum of all the interaction effects between factors F and G, is different from zero.
2nd second-order effect	Do policymakers with different familiarity levels (high or low), and teams (one of the seven teams) differ when considered jointly on the variables demand, supply, understanding and situation awareness?	<ol style="list-style-type: none"> 9. H5a: the sum of all the interaction effects between factors F and T, is equal to zero. 10. H5b: the sum of all the interaction effects between factors F and T, is different from zero.
3rd second-order effect	Do policymakers with different gender (female or male) and teams (one of the seven teams), differ when considered jointly on the variables demand, supply, understanding and situation awareness?	<ol style="list-style-type: none"> 11. H6a: the sum of all the interaction effects between factors G and T, is equal to zero. 12. H6b: the sum of all the interaction effects between factors F and T, is different from zero.
Third-order interaction effect	Do policymakers with different familiarity levels (high or low), gender (female or male) and teams (one of the seven teams), differ when considered jointly on the variables demand, supply, understanding and situation awareness?	<ol style="list-style-type: none"> 13. H7a: the sum of all the interaction effects between factors F, G and T is equal to zero. 14. H7b: the sum of all the interaction effects between factors F, G and T is different from zero.

The analysis starts with the study of the highest order interaction effect (the third order interaction effect) (H7). If the highest order interaction effect is statistically significant, we study the simple second-order interaction effects (the interaction effects of two factors at each level of the third factor, i.e., H4–H6)). If some of the simple second-order interaction effects are significant, we examine the simple main effects (H1–H3). If at least one simple main effect is significant, we compute and interpret the simple comparisons between various factor levels.

If the third-order interaction effect is not significant, we inspect the second-order interaction effects. If some of them are significant, we compute the simple main effects. If none of the second-order interaction effects is significant, we can either finish the analysis or examine the main effects (if they hold interest).

Appendix D. Detailed Factorial MANOVA Results

An ANOVA using Friedman’s test and Tukey’s test for non-additivity for SA scores was conducted. The ANOVA shows that there is a statistically significant increase in situation awareness at the $p < 0.05$ level, $F(1, 34) = 26.85, p = 0.005$. The ANOVA test details are in Table A1 (Appendix A).

Each of the seven Nzoia WeShareIt teams consists of five water policymakers, each representing the five basin county governments (Bungoma, Busia, Kakamega, Trans Nzoia and Uasin Gishu). Since each team has five members, and that SA (and similarly D, A, and U) is measured twice (pre-game and post-game), during the quasi-experiment, we have in total 70 measures of SA. Table A2 (Appendix C) lists effects, questions, and hypotheses for the three IDVs (gender, team, and familiarity) and their interaction effects. Table A3 contains the main descriptive statistics.

Table A3. Descriptive statistics for familiarity factor.

Dependent Variable:			
Familiarity	Mean	Std. Deviation	N
Low	10.1429	7.87828	35
High	26.8571	5.20827	35
Total	18.5000	10.71468	70

Subsequently, 3×4 factorial MANOVA was conducted to compare the effect of three IDVs (gender, familiarity, and team) on the overall situation awareness as well as on the three SA dimensions (demand, supply, and understanding). The between-subject factors can be found in Table A4.

Table A4. Between subject factors.

IDV		Value Label	N
Gender	1	Female	24
	2	Male	46
Familiarity	0	Low	35
	1	High	35
Team	1	Busia*1	10
	2	Busia*2	10
	3	Kakamega*3	10
	4	Bungoma*4	10
	5	Bungoma*5	10
	6	Trans-Nzoia*6	10
	7	Trans-Nzoia*7	10

The factorial analysis started with the study of the highest order interaction effect (the third-order interaction effect), followed by the simple second-order interaction effects (the interaction effects of

two factors at each level of the third factor). Table A5 contains the detailed results of the multivariate tests for the SA dimensions (demand, supply, and understanding) and the overall SA. The results are reported using the Wilks' lambda and Pillai's trace tests. Since we could not conduct Box test of equality of variance, we were not sure whether the assumption of equality of covariances is met. As a consequence, we decided to maintain both the Pillai's trace test (for when the assumption is not met) and the Wilks lambda (when the assumption is met). Otherwise, the Wilks lambda is preferred when the assumption of equality of covariances is met. Both results were relatively similar and the significance levels reported using both tests were the same.

The third order interaction effect assessed whether policymakers with different familiarity levels (high or low), gender (female or male) and teams (one of the seven teams), differ when considered jointly on the variables demand, supply, understanding and situation awareness. There was a significant difference between the levels of familiarity levels (high or low), gender (female or male) and teams (one of the seven teams), differ when considered jointly on the variables demand, supply, understanding and situation awareness, (Wilk's $\Lambda = 2.82$ ($F 9, 112.10$) = 0.78, $p = 0.01$, partial $\eta^2 = 0.15$).

Since the highest order interaction effect is statistically significant, we proceeded to study the simple second-order interaction effects (the interaction effects of two factors at each level of the third factor). Table A5 contains the detailed results of the multivariate tests for the simple second-order interaction effects. The studied three second-order interaction effects, as follows:

1. H4—Familiarity*Gender—1st second-order interaction effect (Do policymakers with different familiarity levels (high or low) and gender (female or male) differ when considered jointly on the variables demand, supply, understanding and situation awareness?)
2. H5—Familiarity*Team—2nd second-order interaction effect (Do policymakers with different familiarity levels (high or low), and teams (one of the seven teams) differ when considered jointly on the variables demand, supply, understanding and situation awareness?)
3. H6—Gender*Team—3rd second-order interaction effect (Do policymakers with different gender (female or male) and teams (one of the seven teams), differ when considered jointly on the variables demand, supply, understanding and situation awareness?)

Table A5. Between subject factors multivariate tests results of gender, familiarity, team factors and their second- and third-order interaction effects on SA and its dimensions (demand, supply, and understanding) for the pre-game and post-game findings.

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Gender	Pillai's Trace	0.05	0.78	3.00	46.00	0.51	0.05
	Wilks' Lambda	0.95	0.78	3.00	46.00	0.51	0.05
Familiarity	Pillai's Trace	0.84	82.74	3.00	46.00	0.00	0.84
	Wilks' Lambda	0.16	82.74	3.00	46.00	0.00	0.84
Team	Pillai's Trace	0.27	0.78	18.00	144.00	0.72	0.09
	Wilks' Lambda	0.75	0.76	18.00	130.59	0.74	0.09
Gender * Familiarity	Pillai's Trace	0.05	0.82	3.00	46.00	0.49	0.05
	Wilks' Lambda	0.95	0.82	3.00	46.00	0.49	0.05
Gender * Team	Pillai's Trace	0.06	0.31	9.00	144.00	0.97	0.02
	Wilks' Lambda	0.94	0.30	9.00	112.10	0.97	0.02
Familiarity * Team	Pillai's Trace	0.44	1.38	18.00	144.00	0.15	0.15
	Wilks' Lambda	0.61	1.37	18.00	130.59	0.16	0.15
Gender * Familiarity * Team	Pillai's Trace	0.41	2.50	9.00	144.00	0.01	0.14
	Wilks' Lambda	0.61	2.82	9.00	112.10	0.01	0.15

As such, we confirm that the null hypotheses H4a, H5a, and H6a are supported:

1. H4a the sum of all the interaction effects between factors F and G is equal to zero;
2. H5a: the sum of all the interaction effects between factors F and T, is equal to zero; and

3. H6a: the sum of all the interaction effects between factors G and T, is equal to zero.

Therefore H4b, H5b, and H6b are not supported:

1. H4b: the sum of all the interaction effects between factors F and G, is different from zero;
2. H5b: the sum of all the interaction effects between factors F and T, is different from zero; and
3. H6b: the sum of all the interaction effects between factors G and T, is different from zero.

Since the highest order interaction effect is significant, but all the three second-order interaction effects are not significant, we proceeded to compute and examine the simple main effects. The main effects analysis contained three main effects:

1. H1—Familiarity—1st main effect (Does high and low familiarity differ on the three subscales and the overall SA scale?)
2. H2—Gender—2nd main effect (Do male and female policymakers differ on the three subscales and the overall SA scale?)
3. H3—Team—3rd main effect (Is there a difference between the seven (7) teams on the three subscales and the overall SA scale?)

For the 1st main effect, a 2×4 MANOVA was conducted with familiarity (low or high) as the independent variable and the overall situation awareness as well as on the three SA dimensions (demand, supply, and understanding) as the dependent variables (Table A5). The results of the MANOVA indicated that there was a significant difference between high and low familiarity on the three subscales and the overall SA, (Wilk's $\Lambda = 0.16$ ($F 3, 46$) = 82.74, $p = 0.005$, partial $\eta^2 = 0.84$).

For the 2nd main effect, a 2×4 MANOVA was conducted with gender (female or male) as the independent variable and the overall situation awareness as well as on the three SA dimensions (demand, supply, and understanding) as the dependent variables (Table A5). The results of the MANOVA indicated that there was no significant difference between female and male SART scores on the three subscales and the overall SA, (Wilk's $\Lambda = 0.95$ ($F 3, 46$) = 0.78, $p = 0.51$, partial $\eta^2 = 0.05$).

For the 3rd main effect, a 7×4 MANOVA was conducted with team (Busia*1, Busia*2, Kakamega*3, Bungoma*4, Bungoma*5, Trans Nzoia*6 and Trans Nzoia*7) as the independent variable and the overall situation awareness as well as on the three SA dimensions (demand, supply and understanding) as the dependent variables (Table A5). The results of the MANOVA indicated that there was no significant difference between high and low familiarity on the three subscales and the overall SA, (Wilk's $\Lambda = 0.75$ ($F 18, 130.59$) = 0.76, $p = 0.74$, partial $\eta^2 = 0.09$).

As such, we confirm that H2a (the group means for the gender (G) factor are equal in the total population); and H3a (the group means for the team (T) factor are equal in the total population), are supported. Therefore, we did not reject these two null hypotheses. Thus, their alternative hypothesis H2b (the group means for the gender (G) factor are different in the total population); and H3b (the group means for the team (T) factor are different in the total population), are not supported. Therefore, we did not reject null hypotheses H2a and H3a.

The 2×4 MANOVA assessment of the familiarity IDV indicates that H1a is not supported. There is a significant difference between the group means for the familiarity (F) factor the overall situation awareness as well as on the three SA dimensions (demand, supply, and understanding) as the dependent variables (Table A5). Therefore, we rejected H1a because the alternative hypothesis H1b is supported (the group means for the familiarity (F) factor are different in the total population). As such, we conclude that the null H2a and H3a are supported, and null H1a is not supported. Alternatively, H1b is supported, and H2b and H3b are not supported.

Table A6. Tests of between-subjects effects.

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Gender	Demand	6.848	1	6.85	0.59	0.44	0.01
	Supply	0.675	1	0.68	0.03	0.86	0.00
	Understanding	18.148	1	18.15	1.75	0.19	0.04
	SA	6.075	1	6.08	0.15	0.70	0.00
Familiarity	Demand	1063.939	1	1063.94	92.27	0.00	0.66
	Supply	3256.546	1	3256.55	148.81	0.00	0.76
	Understanding	1865.295	1	1865.29	180.26	0.00	0.79
	SA	4574.782	1	4574.78	112.96	0.00	0.70
Team	Demand	26.487	6	4.41	0.38	0.89	0.05
	Supply	101.919	6	16.99	0.78	0.59	0.09
	Understanding	63.419	6	10.57	1.02	0.42	0.11
	SA	294.892	6	49.15	1.21	0.32	0.13
Gender * Familiarity	Demand	9.259	1	9.26	0.80	0.37	0.02
	Supply	3.675	1	3.68	0.17	0.68	0.00
	Understanding	6.848	1	6.85	0.66	0.42	0.01
	SA	5.490	1	5.49	0.14	0.71	0.00
Gender * Team	Demand	18.467	3	6.16	0.53	0.66	0.03
	Supply	27.811	3	9.27	0.42	0.74	0.03
	Understanding	2.148	3	0.72	0.07	0.98	0.00
	SA	35.817	3	11.94	0.29	0.83	0.02
Familiarity * Team	Demand	146.689	6	24.45	2.12	0.07	0.21
	Supply	147.395	6	24.57	1.12	0.36	0.12
	Understanding	87.540	6	14.59	1.41	0.23	0.15
	SA	177.321	6	29.55	0.73	0.63	0.08
Gender * Familiarity * Team	Demand	192.748	3	64.25	5.57	0.00	0.26
	Supply	100.745	3	33.58	1.53	0.22	0.09
	Understanding	6.370	3	2.12	0.21	0.89	0.01
	SA	554.557	3	184.85	4.56	0.01	0.22

The tests of Between-Subjects effects provide more insights into the effects of familiarity and the third order interaction effects on SA and its three dimensions (demand, supply, and understanding). For the familiarity factor, when considered alone, there was significant demand effect ($F(1, 6) = 92.27$, $p = 0.005$, partial $\eta^2 = 0.66$), with the post-game SART results reporting significantly higher familiarity on demand for attentional resources than the pre-game SART results.

In summary, the post-game SART results reported significantly higher familiarity effects on all the four dependent variables [overall SA, demand on attentional resources (D), supply of attentional resources (A), and the understanding of the situation that they face at that particular moment (U)], than the pre-game SART results.

Follow up univariate tests of Between-Subjects effects were also conducted for third and second order interaction effects of Gender * Familiarity * Team (third-order effect), Familiarity * Gender (2nd order effect), Familiarity * Team (2nd order effect), and Gender * Team (2nd order effect). The results confirm the MANOVA results. There was no significant second order interaction effect. However, the third order interaction effect indicates mixed results (Table A6).

The post-game SART results reported significantly higher familiarity effects on all the four dependent variables [overall SA, demand on attentional resources (D), the supply of attentional resources (A), and the understanding of the situation that they face at that particular moment (U)] than the pre-game SART results. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

Since one simple main effect is significant, familiarity, we computed a final *simple comparisons* step between low and high familiarity. Since the p -value is lower than 5%, the difference between the factor groups is significant for all the 4 dependent variables in the case of familiarity and significant on only demand and the overall SA on the combined $F*G*T$ third order effect. In other words, familiarity has an essential influence on both demand, supply, understanding and the overall SA. However, the combined effect of $F*G*T$ has an important influence on only demand and the overall SA.

Table A7. Between Subject Factors.

		Dependent Variable:				
(I) Familiarity		Mean Difference (I–J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
					Lower Bound	Upper Bound
Low	High	−16.71 *	1.60	0.000	−19.900	−13.53
High	Low	16.71 *	1.60	0.000	13.529	19.90

Based on estimated marginal means. * The mean difference is significant at the 0.05 level. ^b Adjustment for multiple comparisons: Bonferroni.

Overall SA has an important influence on the dependent variable familiarity. Additionally, D and SA have a significant influence on the third-order combined effect of F*G*T. To assess how big is the influence of SA on the dependent variable familiarity, we assessed the difference between the groups by consulting the table of pairwise comparisons (Table A9, see also Tables A8 and A10).

Table A8. Tests of between-subjects effects for familiarity factor.

Dependent Variable:						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Familiarity	1251.66	1	1251.66	91.16	0.000	0.57
Error	933.71	68	13.73			
Total	13,048.00	70				
Corrected Total	2185.37	69				

Table A9. Pairwise comparisons.

		Dependent Variable:				
(I) Familiarity		Mean Difference (I–J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
					Lower Bound	Upper Bound
Low	High	−16.71 *	1.60	0.000	−19.900	−13.53
High	Low	16.71 *	1.60	0.000	13.529	19.90

Based on estimated marginal means. * The mean difference is significant at the 0.05 level. ^b Adjustment for multiple comparisons: Bonferroni.

Table A10. Univariate tests for familiarity factor.

Dependent Variable:						
	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Contrast	1251.66	1	1251.66	91.16	0.000	0.57
Error	933.71	68	13.73			

The F tests the effect of Familiarity. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

References

1. UN News Centre. COP21: UN Chief Hails New Climate Change Agreement as ‘Monumental Triumph’. *UN News*, 12 December 2015.
2. Victor, D.G. *Global Warming Gridlock: Creating more Effective Strategies for Protecting the Planet*; Cambridge University Press: Cambridge, UK, 2011.
3. United Nations (UN). *United Nations Framework Convention on Climate Change (UNFCCC)*; FCCC/INFORMAL/84; United Nations, Ed.; United Nations: Geneva, Switzerland, 1992; 25p.
4. Kyoto Protocol. *United Nations Framework Convention on Climate Change*; Kyoto Protocol: Kyoto, Japan, 1997; Volume 19.
5. Pearson, B. Market failure: Why the Clean Development Mechanism won’t promote clean development. *J. Clean Prod.* **2007**, *15*, 247–252. [[CrossRef](#)]
6. Falkner, R. The Paris Agreement and the new logic of international climate politics. *Int. Aff.* **2016**, *92*, 1107–1125. [[CrossRef](#)]

7. Falkner, R.; Stephan, H.; Vogler, J. International climate policy after Copenhagen: Towards a ‘building blocks’ approach. *Glob. Policy* **2010**, *1*, 252–262. [[CrossRef](#)]
8. United Nations (UN). *The Paris Agreement*; United Nations, Ed.; United Nations: Paris, France, 2015; p. 27.
9. Change, I.C. *Mitigation of Climate Change*; Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2014.
10. Nachmany, M.; Fankhauser, S.; Townshend, T.; Collins, M.; Landesman, T.; Matthews, A.; Pavese, C.; Rietig, K.; Schleifer, P.; Setzer, J. *The GLOBE Climate Legislation Study: A Review of Climate Change Legislation in 66 Countries*; GLOBE International: Melbourne, Australia, 2014.
11. Giddens, A. *The Politics of Climate Change*; Polity: Cambridge, UK, 2009.
12. Onencan, A.; Enserink, B.; van de Walle, B.; Chelang’a, J. Coupling Nile Basin 2050 Scenarios with the IPCC 2100 Projections for Climate-induced Risk Reduction. *Procedia Eng.* **2016**, *159*, 357–365. [[CrossRef](#)]
13. Onencan, A.; Kortmann, R.; Kulei, F.; Enserink, B. MAFURIKO: Design of Nzoia basin location based flood game. *Procedia Eng.* **2016**, *159*, 133–140. [[CrossRef](#)]
14. Onencan, A.M.; Enserink, B. *THE NILE BASIN BY 2050: Strategic Foresight on the Nile Basin Water Governance*; Nile Basin Discourse: Entebbe, Uganda, 2014; p. 28.
15. Onencan, A.M.; Van de Walle, B. Equitable and Reasonable Utilization: Reconstructing the Nile Basin Water Allocation Dialogue. *Water* **2018**, *10*, 32. [[CrossRef](#)]
16. 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; pp. 43–57.
17. 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, Boston, MA, USA, 7–9 June 2016; pp. 357–365.
18. Onencan, A.; 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; pp. 307–315.
19. Onencan, A.M. TU Delft Serious Game Elevates Nzoia; Flows, 2017. Available online: <https://www.narcis.nl/publication/RecordID/oai:tudelft.nl:uuid%3A91914ae3-453b-458a-ab58-555789a4533c> (accessed on 6 June 2018).
20. Enserink, B.; Onencan, A. Nile Basin Scenario Construction. In Proceedings of the IAIA’s Contribution in Addressing Climate Change, 37th Annual Conference of the International Association for Impact Assessment, Le Centre Sheraton, Montreal, QC, Canada, 4–7 April 2017.
21. Kenya Commission on Revenue Allocation. *Kenya: County Factsheets*; Government of Kenya: Nairobi, Kenya, 2013.
22. Endsley, M.R. Towards a Theory of Situation Awareness in Dynamic Systems. *Hum. Factors* **1995**, *37*, 32–64. [[CrossRef](#)]
23. Wickens, C.D. Situation awareness: Review of Mica Endsley’s 1995 articles on situation awareness theory and measurement. *Hum. Factors* **2008**, *50*, 397–403. [[CrossRef](#)] [[PubMed](#)]
24. Van de Walle, B.; Bruggemans, B.; Comes, T. Improving situation awareness in crisis response teams: An experimental analysis of enriched information and centralized coordination. *Int. J. Hum. Comput. Stud.* **2016**, *95*, 66–79. [[CrossRef](#)]
25. Wright, G.; Goodwin, P. Decision making and planning under low levels of predictability: Enhancing the scenario method. *Int. J. Forecast.* **2009**, *25*, 813–825. [[CrossRef](#)]
26. Onencan, A.M.; Enserink, B.; van de Walle, B. *Game Design Concept Report: Application of the WeShareIt Game Elements in Nzoia River Basin*; Delft University of Technology: Delft, The Netherlands, 2018; p. 53.
27. Government of Kenya (GoK), National Environment Management Authority. *Government of Kenya Second National Communication to the United Nations Framework Convention on Climate Change*; National Environment Management Authority: Nairobi, Kenya, 2015; p. 266.
28. Government of Kenya (GoK). *The Climate Change Act, 2016*; GoK, Ed.; No. 11 of 2016; The Government Printers: Nairobi, Kenya, 2016.
29. Government of Kenya (GoK). *National Climate Change Response Strategy*; MoEaM Resources, Ed.; Government of Kenya: Nairobi, Kenya, 2010.

30. Government of Kenya (GoK). *National Climate Change Action Plan 2013–2017*; MoEaM Resources, Ed.; Government of Kenya: Nairobi, Kenya, 2012; p. 258.
31. Government of Kenya (GoK). *Kenya National Adaptation Plan 2015–2030: Enhanced Climate Resilience towards the Attainment of Vision 2030 and beyond*; MoEaM Resources, Ed.; Ministry of Environment and Natural Resources: Nairobi, Kenya, 2016; p. 68.
32. Intergovernmental Panel on Climate Change (IPCC). *Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories*; IPCC Task Force on National Green House Gas Inventories; IPCC: Geneva, Switzerland, 1996.
33. USAID. Greenhouse Gas Emissions Factsheet: Kenya. In *Climate Links: A Global Knowledge Portal for Climate Change & Development Practitioners*; USAID: Washington, DC, USA, 2017.
34. Climate Watch. Climate Watch. Climate Watch Tools: Country Profiles. In *Country Profiles: Kenya*; Watch, C., Ed.; Climate Watch: Washington, DC, USA, 2018.
35. Adger, W.N. *Fairness in Adaptation to Climate Change*; MIT Press: Cambridge, MA, USA, 2006.
36. USAID. *Greenhouse Gas Emissions in Kenya*; USAID: Washington, DC, USA, 2017; p. 3.
37. Government of Kenya (GoK). *Kenya Climate Smart Agriculture Strategy—2017–2026*; Ministry of Agriculture, Livestock and Fisheries in Collaboration with the Ministry of Environment and Natural Resources with Support from the World Bank, the United Nations Development Programme (UNDP) and the Food and Agriculture Organization of the United Nations (FAO) through the Integrating Agriculture in National Adaptation Plans Project; Government of the Republic of Kenya: Nairobi, Kenya, 2017; p. 53.
38. Government of Kenya (GoK). *Bus Rapid Transit (BRT) Plus System for the Nairobi Metropolitan Region*; Ministry of Environment, Natural Resources and Regional Development Authorities: Nairobi, Kenya, 2015.
39. Akwei, I. Africa is splitting into two after tear in Kenya's Rift Valley. Narok County, Kenya. *Fece News*, 22 March 2018.
40. Kahongeh, J.; Macharia, M. Kenya is Splitting at Suswa. *Daily Nation*, 9 April 2018.
41. Skilling, I. Incremental caldera collapse of Suswa volcano, Gregory rift valley, Kenya. *J. Geol. Soc.* **1993**, *150*, 885–896. [[CrossRef](#)]
42. Biggs, J.; Anthony, E.; Ebinger, C. Multiple inflation and deflation events at Kenyan volcanoes, East African Rift. *Geology* **2009**, *37*, 979–982. [[CrossRef](#)]
43. Mayer, I.; Bekebrede, G.; Harteveld, C.; Warmelink, H.; Zhou, Q.; Van Ruijven, T.; Lo, J.; Kortmann, R.; Wenzler, I. The research and evaluation of serious games: Toward a comprehensive methodology. *Br. J. Educ. Technol.* **2014**, *45*, 502–527. [[CrossRef](#)]
44. Mayer, I.; Bekebrede, G.; Warmelink, H.; Zhou, Q. A brief methodology for researching and evaluating serious games and game-based learning. In *Psychology, Pedagogy, and Assessment in Serious Games*; IGI Global: Hershey, PA, USA, 2013; pp. 357–393.
45. Mayer, R.C.; Davis, J.H.; Schoorman, F.D. An integrative model of organizational trust. *Acad. Manag. Rev.* **1995**, *20*, 709–734. [[CrossRef](#)]
46. Luhmann, N. Familiarity, confidence, trust: Problems and alternatives. *Trust Mak. Break. Cooper. Relat.* **2000**, *6*, 94–107.
47. Mayer, I.; Veeneman, W.; The, D. *Games in a World of Infrastructures*; SimulationGames for Research Learning and Intervention; Eburon: Delft, The Netherlands, 2002.
48. Zsombok, C.E.; Klein, G. *Naturalistic Decision Making*; Psychology Press: New York, NY, USA, 2014.
49. Klein, G.A.; Orasanu, J.E.; Calderwood, R.E.; Zsombok, C.E. Decision Making in Action: Models and Methods. In *Proceedings of the Workshop*, Dayton, OH, USA, 25–27 September 1989; Ablex Publishing: New York, NY, USA, 1993.
50. Klein, G.A. *A Recognition-Primed Decision (RPD) Model of Rapid Decision Making*; Ablex Publishing Corporation: New York, NY, USA, 1993.
51. Garriss, R.; Ahlers, R.; Driskell, J.E. Games, motivation, and learning: A research and practice model. *Simul. Gaming* **2002**, *33*, 441–467. [[CrossRef](#)]
52. Landers, R.N. Developing a theory of gamified learning: Linking serious games and gamification of learning. *Simul. Gaming* **2014**, *45*, 752–768. [[CrossRef](#)]
53. Onencan, A.M.; Enserink, B.; Van De Walle, B. Nzoia WeShareIt Situation Awareness Dataset. In *Nzoia WeShareIt Research Program*; Onencan, A.M., Ed.; 4TU: Delft, The Netherlands, 2018.
54. Taylor, R. Situational Awareness Rating Technique(SART): The development of a tool for aircrew systems design. In *Proceedings of the Aerospace Medical Panel Symposium*, Copenhagen, Denmark, 2–6 October 1989.

55. 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: Berlin, Germany, 2017; pp. 161–172.
56. Endsley, M.R. Measurement of situation awareness in dynamic systems. *Hum. Factors* **1995**, *37*, 65–84. [[CrossRef](#)]
57. Sarter, N.B.; Woods, D.D. Situation awareness: A critical but ill-defined phenomenon. *Int. J. Aviat. Psychol.* **1991**, *1*, 45–57. [[CrossRef](#)]
58. Mathews, J.A. Naturalizing capitalism: The next Great Transformation. *Futures* **2011**, *43*, 868–879. [[CrossRef](#)]
59. 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](#)]
60. Shubik, M. *Games for Society, Business, and War*; Elsevier: New York, NY, USA, 1975.
61. 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.
62. Mayer, I.S. The gaming of policy and the politics of gaming: A review. *Simul. Gaming* **2009**, *40*, 825–862. [[CrossRef](#)]
63. Stone, D. *Policy Paradox*, rev. ed.; WW Norton & Company: New York, NY, USA, 2002.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).