Sustainability indicators

Monitoring cross-county water cooperation in the Nzoia river basin, Kenya

Onencan, Abby Muricho; Enserink, Bert; Van de Walle, Bartel

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Sustainability Indicators: Monitoring Cross-County Water Cooperation in the Nzoia River Basin, Kenya

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

Policy Analysis Section, Multi-Actor Systems (MAS) Department, Faculty of Technology,
Policy and Management, Delft University of Technology, Building 31, Jaffalaan 5, 2628 BX Delft, P.O. Box 5015,
2600 GA Delft, The Netherlands; B.Enserink@tudelft.nl (B.E.); B.A.vandeWalle@tudelft.nl (B.V.d.W.)
* Correspondence: a.m.onencan@tudelft.nl; Tel.: +31-15-27-81-810

Abstract: Kenya Water Services Regulatory Board (WASREB) Impact Report indicates a stagnation in water coverage at 55 percent, for the last three years, contrary to the 2015 target of 80 percent. One main reason for the stagnation is weak cross-county cooperation between hydrologically interdependent governments. WASREB has little guidance on what indicators to use to enhance cross-county water cooperation. Through literature review, we assess whether the UN-Water methodology for assessing Sustainable Development Goals (SDG) 6.5.2 would provide useful guidelines. Based on the literature review outcomes, we design a water policy game known as Nzoia WeShareIt. After that, we play seven-game sessions in four county governments (Busia, Bungoma, Kakamega, and Trans Nzoia), on 11–22 July 2016. We use the in-game and post-game questionnaire data to measure learning outcomes on interdependence and cooperation. The findings indicate that Nzoia WeShareIt policy game as a form of experiential learning increased understanding on the value of cross-county cooperation. The study constitutes a practical guideline to WASREB and a quick reference tool to be explored when designing indicators to monitor cross-county cooperation. We also propose a mixed method approach that incorporates team interdependence indicators as distinct and separate indicators from cooperation. Moreover, we recommend strengthening SDG 6.5.2 indicator to measure transboundary water cooperation inputs, processes and outcomes.

Keywords: Sustainable Development Goals (SDG); SDG Indicator 6.5.2.; transboundary water cooperation; team interdependence; water policy games; Nzoia river basin; sustainability; equity; benefit sharing

1. Introduction

The Kenyan National Water Service Strategy (NWSS) 2015 target was 80 percent access to safe water in urban centres and 75 percent in rural centres [1] (p. 7). The Water Services Regulatory Board (WASREB) 2018 Impact Report, documents that water coverage stagnated at 53 percent for three years (2011–2014) and later at 55 percent for another three years (2014–2017) [2] (p. 19). The 2009 population census reports 28 percent access to piped water and 37 percent to non-piped water (springs, wells, and boreholes) [3] (p. 22).

Kenya water resources are sufficient to meet the 2030 projected demand, thus the current water scarcity is owing to governance failure [4,5]. Global Water Partnership (GWP) defines water governance as "the range of political, social, economic and administrative systems that are in place to develop and manage water resources, and deliver water services, at different levels of society" [6]. According to the 2018 Water Resources Authority (WRA) Situation Report, the 2030 projected demand is 21,468 Million Cubic Meters (MCM)/year, against available water resources of 26,634 MCM/year.

The WRA situation report identifies three main governance challenges that need to be addressed to resolve water scarcity in Kenya. First, there is a significant water demand variance from various geographical regions (county governments). Second, the geographical distribution of available water varies considerably [7]. Thirdly, there is a considerable strain on stored water resources, especially during the dry spells, due to low water storage capacity, population pressure, rapid urbanization, anthropogenic activities, and climate change. In the last eight years, the population has significantly increased from 39,799,151 to 51,571,283 (population on 28th of December 2018) [8]. Addressing the three core governance problems will advance Sustainable Development Goal (SDG) 6 on ensuring "availability and sustainable management of water and sanitation for all" Kenyans [9] (p. 9).

WASREB is charged with the responsibility of monitoring the implementation of SDG 6 [9]. WASREB was re-constituted under the 2016 Water Act [10], and its mandate enhanced from water regulation to monitoring and issuing water licenses [2,3]. Figure 1 represents the Kenya institutional framework after the enactment Water Act 2016.
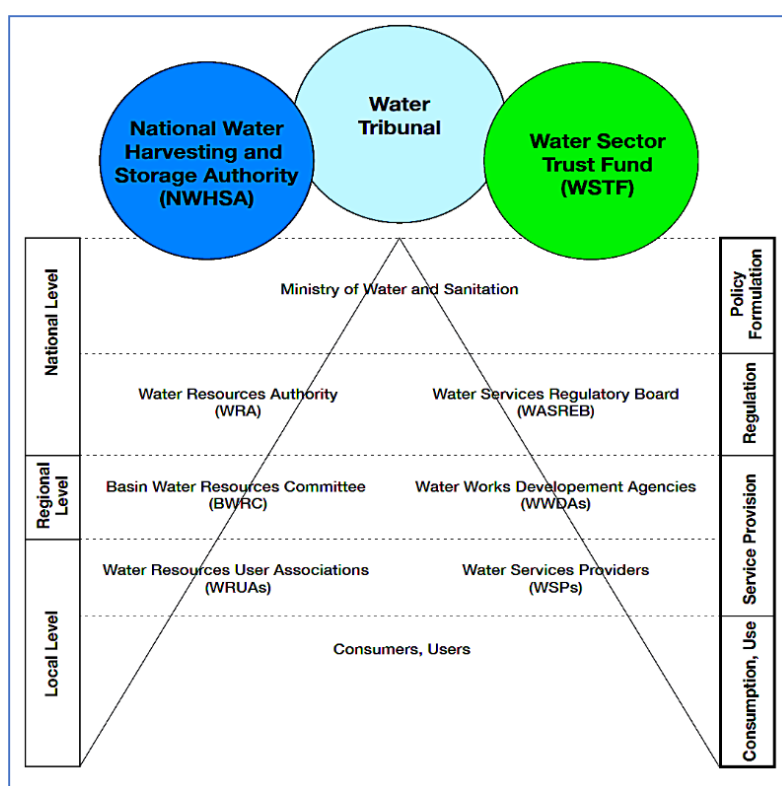

Figure 1. Framework as outlined in the Water Act 2016. Source: Kenya Water Services Regulatory Board (WASREB) [2] (p. 12).

County governments are responsible for ensuring that Kenyan residents within their county have access to safe water and sanitation. They exercise their responsibility through establishing public limited liability companies, known as Water Service Providers (WSPs) [3] (p. 17). WSPs can either be within a county or be cross-county. We define cooperation as the voluntary act by two or more riparian governments to jointly engage in an exchange which benefits all parties through the sharing of river basin water resources, and the creation of new resources or both [11]. Cross-county cooperative arrangements are managed by the eight national Water Service Boards (Boards).

The study focus is Nzoia River Basin, in western Kenya. It is a sub-basin of Lake Victoria Basin. Nzoia river transverses through six county governments. Two downstream counties (Busia and Siaya), two middle stream counties (Bungoma and Kakamega) and two upstream counties (Uasin Gishu and Trans Nzoia). The Lake Victoria North WSB management of the Nzoia River Basin is piecemeal.

Siaya, a downstream riparian government is not part of the WSB. Moreover, the WSB manages county governments outside Nzoia Basin (Keiyo Marakwet, Nandi, and Vihiga) [3].

Nzoia River Basin county governments have no agreement or formal arrangement for the joint management of the basin. However, through the Nzoia Water Services Company Limited (NZOWASCO) and the Lake Victoria Basin Commission (LVBC), there is evidence of cooperation between Bungoma, and Trans Nzoia. NZOWASCO delivers water services within its framework of cluster towns. The cluster towns approach only applies to the urban centers in the local government, thus leaving a large area of the local government, which is mainly rural. LVBC covers all the Lake Victoria Basin countries and county governments [12–14]. Therefore, its scope is much broader and not customized for the Nzoia River Basin.

WASREB is the only institution that monitors county government performance through WSPs. County governments are the sole institution constitutionally responsible for the provision of water and sanitation services. Particularly, they are absent in the 2016 Water Act institutional framework (Figure 1), [3]. Therefore, it is difficult to hold them accountable for responsibilities beyond the scope of WSPs. WASREB plays a critical role of monitoring county government actions and documenting the outcomes in the Annual Impact Reports, to increase transparency and accountability [2].

The 2016/2017 WASREB Impact Report raises some critical water governance issues that need to be addressed [3] (pp. 29–37). Some of these critical issues are:

1. Devolution as provided in the 2010 constitution [15] and subsequently in the 2016 Water Act [10] has led to conflicts between the central government and the county governments leading to vexatious litigations, stalling of planned water projects and subsequently a decline in water service provision [3] (p. 31);
2. The commercial sector is encroaching catchment areas, risking water resources [3] (p. 30);
3. Citizens are polluting the river water with waste and straining service provision [3] (p. 30);
4. Lack of adequate sewerage infrastructure has led to more river pollution, thus increasing water treatment costs [3] (p. 31);
5. Climate change is not aligned with current efforts to conserve water, thus threatening the sustainability of water towers [3] (p. 31); and
6. Poor land practices have led to degradation and siltation of the dams [3] (p. 31).

To address some of these governance challenges, WASREB monitors WSPs using ten Key Performance Indicators (KPIs), namely: (1) Water Coverage; (2) Drinking Water Quality; (3) Hours of Supply; (4) Non-Revenue Water; (5) Metering Ratio; (6) Staff Productivity Measured by Staff per 1000 Connections; (7) Personnel Expenditure; (8) Revenue Collection Efficiency; (9) Operation and Maintenance Cost Coverage; and (10) Sewerage. None of the indicators measures cross-county cooperation nor addresses the key governance challenges facing the water sector, in Kenya [2].

Consequently, WASREB faces two significant challenges. First, monitoring using the KPIs will not resolve the core governance problems as outlined in their environmental scan [2,3]. The governance challenges need to be resolved to create an enabling environment for the WSPs to operate [16–25]. Second, there is a mismatch between the WASREB KPIs and the SDG 6 goal. KPIs enhance competition, unilateral WSP actions, and ultimately inequitable water allocation [25]. Competition and unilateral actions do not promote national access and sustainable water use [18,21,26]. The top ten WSPs are from water-rich counties (Nyeri, Meru, Thika, Nakuru, Ngagaka, Nanyuki, Ngandori Nginda, Malindi, and Kakamega). The bottom ten utilities are mainly from water-scarce arid and semi-arid zones (Lodwar, Tililbei, Kwale, Kitui, Bomet, Wajir, Garissa, Eldama Ravine, and Olkejuado) [2] (p. 21). Migori though not in the arid and semi-arid zone is in the bottom ten WSPs because the resident's taps were dry for five months due to failure by the WSP to pay electricity expenses worth five million Kenya Shillings (Kshs). These energy expenses were incurred when extracting water from the river and distributing the water to residents [27].

The WASREB KPIs are not comparable across the WSPs due to inequitable water distribution, water demand and topography. First, performance of water-rich counties cannot be compared with performance of water scarce counties, with no water to distribute during the dry seasons. Second, performance of cities with dense population, few water sources and residents who rely 100 percent on piped water is not comparable with rural areas. Third, performance of highland counties using gravity method to distribute water cannot be compared with lowland counties that have colossal energy expenses due to extracting the water from the river and later distribute it to the residents. Moreover, if the playing field is not leveled to ensure equity, the water resource-rich counties will continue amassing capital and constructing water storage facilities to perform better. Whereas the water-scarce regions will not receive any revenue from their residents due to their inability to deliver services. Therefore, the gap between the top performers and the bottom performers will continue to widen. This may lead to either continued stagnation or declined national performance of SDG 6.

We undertook this study to support WASREB in monitoring SDG 6. The specific target that we will focus on is 6.5 "By 2030, implement integrated water resources management at all levels, including through transboundary cooperation." Transboundary waters refer to ground and surface waters that "mark, cross, or are located on international political boundaries between two or more States [28]." The Indicator we intend to improve is 6.5.2, on the "proportion of transboundary basin area with an operational arrangement for water cooperation" [29]. Nzoia Basin does not cross international boundaries. Therefore, it is not a transboundary basin under indicator 6.5.2. However, being a sub-basin of Lake Victoria, and Lake Victoria being a sub-basin of the Nile Basin, the results of the studies can be applied to the two larger basins and locally [3].

We first conduct a literature review to identify gaps in implementing SDG 6.5.2 indicator. After that, we design a water policy game to enhance learning on water cooperation [30]. Gaming is an experiential learning method that has proven to be effective in enhancing cooperation through facilitated interactions and shared player experience [31]. Water policy games have the potential to bridge collaboration gaps and ultimately increase interdependence of groups that are geographically dispersed [32,33]. Research on the effect of gaming on team interdependence and cooperation in river basin management remains unexplored [16,17]. Thus, there is little guidance to WASREB on innovations to enhance learning on water cooperation. Also, at the policy level, there is little guidance on what indicators WASREB should use to measure cross-county water cooperation [33]. With a primary focus on interdependence and cooperation, we designed a game known as Nzoia WeShareIt. After that, we played seven sessions with Nzoia basin water policy makers [34]. We collected in-game and post-game data. We adopted the main recommendations from the Depping and Mandry [33], namely; we incorporated competition as a positive element that enhances collaborative play and assessed cooperation and interdependence as two distinct mechanics.

To address identified science, policy and practice gaps, we seek to answer these questions:

1. Contribution to Science: What are the identified gaps/challenges WASREB may face when implementing the "step-by-step monitoring methodology for Indicator 6.5.2"?
2. Contribution to Practice: Did the WeShareIt game increase learning on team interdependence and cooperation in the Nzoia River Basin?
3. Contribution to Policy: What are the policy recommendations to the Kenyan Water Services Regulatory Board (WASREB) on improving the monitoring of cross-county water cooperation?

The paper is organized as follows. The next section identifies gaps/challenges WASREB may face when implementing the "step-by-step monitoring methodology for Indicator 6.5.2" (Section 2). The third section introduces the conceptual framework for the policy game, details of the Nzoia WeShareIt policy game, subscales, measures, procedure, and data analyses. The fourth section presents the results obtained from the post-game data. The data were analyzed using two methods: Chi-Square test for goodness-of-fit and one-way Analysis of Variance (ANOVA). Then we analyze the results

and provide recommendations to improve the identified science, practice, and policy gaps (Section 5), and provide concluding remarks.

2. Local and Global Gaps to Monitoring Water Cooperation in a Shared River Basin

2.1. Applicability of SDG Indicator 6.5.2 in the Nzoia River Basin in Kenya

UN-Water developed a "step-by-step monitoring methodology for Indicator 6.5.2" [29,35,36]. The methodology defines water cooperation arrangement as treaties (bi/multilateral), agreements, conventions or other formal arrangements (for instance a Memorandum of Understanding between the riparian states) [29]. The definition limits the scope of water cooperation to state actors. It is only State actors that are qualified to enter into such agreements or arrangements, either as individual riparian states, interstates, inter-ministerial, regional agencies/authorities, and inter-governmental bodies [29,37]. The methodology, further provides that the "arrangement for water cooperation" must meet the following set of criteria to be considered "operational":

1. "There is a joint body, joint mechanism, or commission (e.g., a river basin organization) for transboundary cooperation;
2. There are regular (at least once per year) formal communications between riparian countries in form of meetings (either at the political or technical level);
3. There is a joint or coordinated water management plan(s), or joint objectives have been set;
4. There is a regular exchange (at least once per year) of data and information" [29] (p. 3).

Based on the SDG 6.5.2 requirements, Nzoia basin is at 0 percent cooperation. It does not possess:

- A joint basin management institution for cross-county cooperation;
- A basin management plan to jointly and sustainably manage the shared resource [38] (p. 12).
- A data sharing protocol [22], thus water managers hardly possess information on the current water use and quality within the drainage basin [38] (p. 13).
- An information management system [38] (p. 13) on the basin's current, planned and potential future water uses [26], the effects of the rapid population and economic growth [38] (p. 13); and
- A history of strong cooperation [38] (p. 12). There is no institutional framework at the basin level to convene and facilitate basin meetings. Water management is an internal WSP matter.

The Water Resources Authority (WRA) National Water Master Plan 2030 [39] (p. 46) indicate plans to construct several dams along the Nzoia river for irrigation and stop flooding. Flooding will be controlled by setting an environmental flow rate and environmental monitoring. Conversely, SDG Indicator 6.5.2 does not measure steps towards establishing cooperation. Some researchers indicate that the proposed methodology fails to capture and monitor the true state of implementation of Transboundary Water Cooperation (TWC) [37,40]. In the next sub-section, we will assess the identified gaps when implementing Indicator *SDG 6.5.2*, in other river basins.

2.2. Identified Policy Gaps: SDG 6.5.2

Studies propose three methodologies to measure cooperation (Figure 2). The first assesses formal agreements, and is supported by most Inter-Governmental Organizations. This approach, partially measures cooperation outcomes and fails to measure process and input elements. SDG indicator 6.5.2 takes the first approach [37,40]. To strengthen SDG 6.5.2 various researchers propose either a cooperation continuum [41] or qualitative analyses, including hydropolitical assessments and discourse analyses [37,42–55]. Based on relevant literature review, we identified ten Indicator 6.5.2 gaps, as discussed in this sub-section.
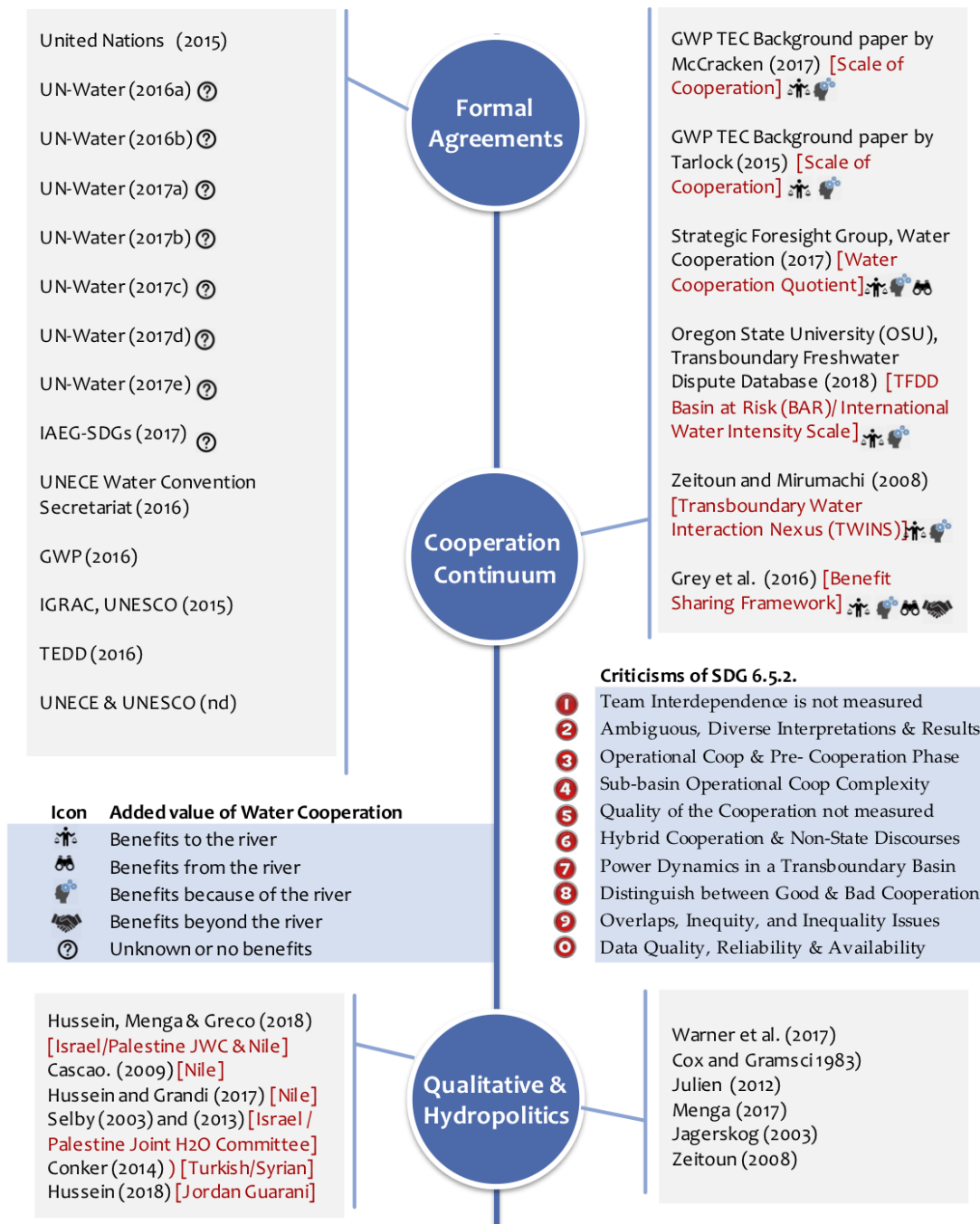
Figure 2. Review on the Current and Proposed Monitoring Instruments for SDG 6.5.2.

2.2.1. Gap 1: Team Interdependence is not Measured as a Distinct Indicator from Cooperation

SDG 6.5.2 measures the outcomes of cooperation with negligible attempts to measure the process and no provision to measure input elements. Team interdependence is input element, which if lacking, the process and outcomes may not materialise. Research indicates that interdependence is the most critical element for team formation [30] (p. 201). A team is a group of persons that interact adaptively and interdependently, to achieve their "specified, shared and valued" objective(s) [56] (p. 3). Studies confirm that hydrological interdependence does not automatically translate into team formation and cooperation [16–22,26,57,58]. River basin groups lack the basic elements of team formation, namely,

a common purpose and individuals who "interact adaptively and independently to achieve specified, shared and valued objectives" [56] (p. 3). The water sector is lagging in developing strong river basin teams that work interdependently to achieve a shared goal [38,59].

Lack of team interdependence negatively affects the quality of cooperation [24,33]. Hall (2014), argues that without interdependence, many groups proceed with their planned unilateral actions, without any form of cooperation and never mature into a team [30] (p. 201). If the river basin group does not mature into a team, then they cannot cooperate [38]. To facilitate cooperation, there is a need for capacity development of WSPs in team interdependence [32,33].

2.2.2. Gap 2: SDG 6.5.2. Is Ambiguous, Leading to Diverse Interpretations and Results [40] (p. 9)

SDG 6.5.2 criteria to measure operationality of cooperation is ambiguous [37,60]. It is comprised of four components, namely: (1) joint river basin management plans; (2) joint river basin institution/organization; (3) meetings; and (4) data exchange [61]. According to a study by McCracken and Meyer [40] (p. 9), most continents got zero percent, at the transboundary aquifer level (zero percent for America, Asia, and the Middle East and 0.1 percent for Europe). At the river basin level, America and Africa were leading (90.8% and 67.1%, respectively), while Europe (31.7%), Asia and the Middle East (11.7%) were still lagging. Different countries and continents defined the attributes differently leading to results that cannot be compared. Meetings and the joint river basin management were construed widely or extremely narrowly, leading to diverse results [40].

2.2.3. Gap 3: Operational Indicator Masks Pre-Cooperation Phase [40] (p. 9)

SDG 6.5.2. fails to measure the pre-cooperation phase [37]. McCracken and Meyer [40] (p. 9) explain that the SDG 6.5.2 operational cooperation fails to capture the stage at which a riparian state is in the cooperation continuum. It is limited to official/formal cooperation thus leaving out the pre-cooperation phase and stepwise cooperative processes. de Chaisemartin [60] (p. 20) state that "the fact that an arrangement 'operates' as per indicator 6.5.2 does not indicate the quality of the operationality." SDG 6.5.2, fails to capture hybrid forms of cooperation, such as water policy networks, that have become more attractive to state actors in the recent past compared to formal cooperation [17–20,26,58]. To address this gap, Hussein et al. [37] recommend the revision of Indicator 6.5.2 to incorporate formal, informal and technical state deliberations (preoperational phase).

McCracken and Meyer [40] (p. 9) further explain that the binary nature of operational cooperation requires that a riparian state meets all the four criteria, to qualify. Therefore, if a state meets only three of the four criteria, they receive a zero percent rating. The binary nature of the indicator fails to take account of the process and the steps towards cooperation and ranks riparians on two sides of the spectrum. To illustrate this challenge McCracken and Meyer [40] (p. 9) provide an example of the Ganges-Brahmaputra-Meghna basin where the riparian states met three criteria and failed to fulfill the joint basin management criteria. As a consequence, the achieved step-wise cooperative efforts and evolving collaborative actions were not included in the final assessment. It is not clear how to address cooperative actions and arrangements that do not meet the four criteria of operational cooperation.

2.2.4. Gap 4: Sub-Basin Operational Cooperation Complexity (Scale Issues)

The UN-Water 2017 Step-by-Step Methodology for Indicator 6.5.2 [35] was updated (from the April 2016 draft [36]) to incorporate sub-basin scale operational cooperation [40]. Thus, cooperating sub-basins can measure evolving cooperation actions separate from the umbrella basin. Swain [62] explain the difficulty of cooperation for large basins, and recommends that joint cooperative actions should take place at the sub-basin level.

McCracken and Meyer [40] (p. 10) state that the UN-Water 2017 revision to the Step-by-Step Methodology for Indicator 6.5.2, is a positive move towards addressing scale issues, but still leaves two gaps unaddressed. First, there are cooperation actions and arrangements that do not cover the entire basin area or sub-basin. The 1959 agreement between Egypt and Sudan that allocates the Nile

river flows between two countries is a cooperative agreement that covers only a part of the Eastern Nile sub-basin and only two out of the 11 countries [18]. This unique instance is not anticipated nor provision made under SDG 6.5.2 [40] (p. 10). Second, the sub-basin results of SDG 6.5.2 are not comparable. The indicator fails to assess the substantive issues, which make amalgamation of the results at different governance levels inaccurate and unreliable.

2.2.5. Gap 5: Quality of the Cooperation Not Measured

The indicator measures only the outcome indicators, thus excluding process and input indicators [23], and thereby failing to measure the quality of cooperation [37]. McCracken and Meyer [40] explain the deception behind the term "cooperation" and the ongoing hydro-politics that cannot be measured by an agreement. Research indicate that power plays a critical role in the allocation of water resources, which is barely captured through counting the number of agreements a country has signed and ratified [42,50,55]. Hussein et al. [37] recommend the revision of Indicator 6.5.2 to qualitative indicators to assess the quality of cooperation. de Chaisemartin [60], recommend the use of indices to measure the quality of cooperation.

2.2.6. Gap 6: Does Not Acknowledge Hybrid Cooperation and Non-State Actors Discourses

Another criticism of indicator 6.5.2 is its inability to incorporate hybrid forms of cooperation. These networks lie between the formal kind of cooperation where the countries have to sign an agreement and establish a joint body and ad-hoc/fluid types of cooperation. These informal arrangements cannot be measured under indicator 6.5.2 [37]. Moreover, the indicator does not measure the numerous actions by politicians (political statements), media (print, radio, television, social media), non-profit organizations (NGOs), civil societies, and development partners including the World Bank. There is no analysis of the ongoing discourse between different state and non-state actors with competing interests and diverse objectives. Hussein et al. [37] recommend the revision of Indicator 6.5.2 to include qualitative indicators to assess state and non-state discourses and hybrid forms of cooperation where different stakeholders are involved.

2.2.7. Gap 7: Fails to Take Account of Power Dynamics in a Transboundary Basin

SDG 6.5.2 fails to take account of the power dynamics that exist in some river basins. To explain the importance of power dynamics Hussein et al. [37] provides an example of the 1995 Israel—Palestine Liberation Organization Oslo II Agreement. The Oslo treaty "codified and cemented asymmetric power relations and a non-equitable share of water resources between the two parties" [37] (p. 4). Within the so-called cooperation, all Palestinian water development requests were not approved by Israel between 1995 and 2008. Palestinians approved Israeli's requests to construct water supply facilities in the West Bank. Zeitoun (2008) cautions against agreements that seem to portray cooperation, and in reality, they foster domination [63]. Hussein et al. [37] recommend developing indicators that pierce through the formalities and diplomatic subtleties of formal agreements and unveil concealed power asymmetries and domination.

2.2.8. Gap 8: Does Not Distinguish between Good and Bad Cooperation

Indicator 6.5.2 does not make a distinction between good and bad cooperation. Hussein et al. [37] provide instances where transboundary river basins with agreements or formal arrangements experience weak or no cooperation, whereas others have no agreement and enjoy quality cooperation. Indicator 6.5.2 assumes that the presence of an arrangement for cooperation is equivalent to good cooperation. The cooperation between 1959 Egypt and Sudan agreement led to the allocation of the Nile water flows amongst two countries to the exclusion of the upper riparians. Onencan [18] explains the inequity in the 1959 agreement's water allocation. Moreover, Egypt got the largest share of the Nile water flows while Sudan paid considerable to implement the agreement provisions [18].

2.2.9. Gap 9: Overlaps, Inequity, and Inequality Issues

There are many overlaps at the basin and sub-basin level leading to double reporting and distortion of the overall picture. For instance, Democratic Republic of Congo is a Nile River Basin riparian states, and at the same time a riparian state in Congo Basin and Tanganyika Lake sub-basin. The freshwater from Lake Tanganyika flows into River Congo and finally into the Atlantic Ocean. McCracken and Meyer [40], explain that under SDG 6.5.2, a transboundary basin can only be accounted for once. Therefore there is an overlap between the Congo Basin and Tanganyika Lake sub-basin. It is not clear whether the cooperative actions and arrangements of the Democratic Republic of Congo should be attributed to Congo Basin or Tanganyika Lake sub-basin.

Moreover, transboundary aquifers do not lie within specific river basins leading to further overlaps. The overlaps deepen the complexity of monitoring transboundary water cooperation. A single jacket approach to address existing complexities has led to inequitable outcomes [18].

The area as a unit of analysis obscures the real basin situation. To address the complexities that arise from using basin area as a unit of analysis, de Chaisemartin [60], recommend the use of other units of measurement like the area of the riparian state in the basin and the volume of water. McCracken and Meyer [40] (p. 10) propose a shift from the area as the unit of analysis towards "volume of water, number of people dependent on the resource or number of agreements." Area as a unit of analysis provide false impressions that larger basins are more important. In the case of the Nile River, it is the longest river in the world and transverses 11 countries. The basin area is massive, but the volume of the Nile river flows does not correlate with the vast basin area. Also, the river flow is seasonal, most of the water is lost through evaporation in Lake Victoria and the Sudd swamp. Furthermore, 86 percent of the main Nile water is from one country, Ethiopia [64].

2.2.10. Gap 10: Data Quality, Reliability, and Availability Issues

The binary (yes or no) nature of the SDG 6.5.2 indicator makes it difficult to grasp the actual situation from the aggregated data. The pressure to have a yes as opposed to a no may lead to unreliable data, thus affecting the quality. Moreover, there is no mechanism to check the quality of the data which affects the reliability of the data. Some of the indicators are ambiguous and depend on one's interpretation, leading to multiple interpretations and collection of data that is not comparable nor interoperable. Lack of disaggregation of the data makes it difficult to assess the actual situation [40] (p. 9).

3. Materials and Methods

3.1. Nzoia WeShareIt Conceptual Framework

To address the identified indicator 6.5.2 gaps, we designed the Nzoia WeShareit conceptual framework for the Water Policy Game. Table 1 provides how we addressed all the gaps within the Nzoia WeShareIt conceptual framework.

The Nzoia WeShareIt conceptual framework was designed using the serious game input-process-output model by Garris, Ahlers, and Driskell [65]. The model facilitates the measuring of cross-county cooperation at the input level (team interdependence and climate change induced cooperation), process level (steps made towards cooperation and the pre-cooperation phase) and the outcome level (water policy network/formal legal agreement/joint basin management institution). Figure 3 illustrates how the input-process-model was assimilated in the game.

Table 1. Nzoia Game Innovations to Address the SDG 6.5.2 Policy Gaps.

SDG Indicator 6.5.2 Gap	The Contribution of Nzoia WeShareIt Water Policy Game
1. Team Interdependence	• Two sub-scales were developed to measure cooperation and team interdependence separately with separate game mechanics
2. Ambiguous, Diverse Interpretations and Results	• In-game assessment indicators were clear, precise and disaggregated. • The game outcomes were Real-time and displayed on the Whiteboard. • In-game questionnaire in Swahili and discussed before it was applied
3. Operational Cooperation and Pre-Cooperation Phase	• The game measured pre-game, in-game and post-game cooperation outcomes, including operational cooperation. • Water policy networks at the same level as a formal agreement.
4. Sub-basin Operational Coop Complexity	• Each player had different indicators based on the county's circumstances. • Cooperation arrangements that do not cover the basin area were monitored. • Substantive issues related to cooperative arrangements measured.
5. Cooperation Quality is not measured	• Nzoia WeShareIt was designed to assess the input, process, and outcome in the entire policy process (see Figure 3). • Mixed methods were employed to collect and analyze the data. • The in-game, post-game and pre-game data formed the quantitative data. • The post-game questionnaire, debriefing sessions with the policymakers and the rough-cut videos for the game sessions constituted the qualitative data.
6. Hybrid Cooperation & Non-State Discourses	• Non-state policymakers were also invited to play the game. • The discourse between the various actors was captured through mixed method data collection methods and an in-game peer review mechanism. • Hybrid policy networks and measured using in-game trading data [19,25]
7. Power Dynamics in a Transboundary Basin	• The game rules were flexible for the stronger county governments to exercise hydro-hegemony and this was measured through the in-game data & video.
8. Distinguish between Good & Bad Cooperation	• In-game peer review questionnaire filled at the end of every round. • Anonymous results displayed Realtime on a leaderboard. • Discussion of the results from the Real-time feedback and provide an opportunity to improve cooperation.
9. Overlaps, Inequity, and Inequality Issues	• The area is not the unit of analysis. The unit was available water volumes and productivity levels for food and energy production.
10. Data Quality, Reliability & Availability	• Complementary knowledge game mechanic encouraged data sharing. • Shared information was vetted through the in-game peer review mechanism. • Negative reviews for poor quality, unreliable information or for not sharing.

The Nzoia River policymakers make decisions within a deeply uncertain environment occasioned by climate change. They allocate their water resources for food and energy under uncertain future events. Climate change was a crucial element in the conceptual framework, which we utilized to promote cooperative actions. Since climate change disasters are slow-onset, their effects are delayed, and can be catastrophic, if not well planned for [16–21]. We integrated the naturalistic decision-making model to support Nzoia policymakers to make better decisions in a profoundly complex and uncertain environment, that is continuously changing, there is time stress (when the disaster occurs) and the stakes are high (Figure 3). Therefore we adopted the Recognition-Primed Decision (RPD) model by Klein [66] to support policymakers in systematic decision-making, through three phases. First, they perceive the elements in the situation, if the climate change situation is familiar. In the second phase, they review the available options and whether they would work. The last phase they implement the selected option, and enter into binding relationships with other riparian states or establish a network of buyers and sellers. The policymakers have no time to assess documents before they make the decisions, therefore their policy actions are selected from cues based on past experience and lessons learnt during the game sessions. This approach helps overcome indecision that normally occurs when policymakers are faced with an unfamiliar world with no past knowledge to inform present and future actions [19,25].
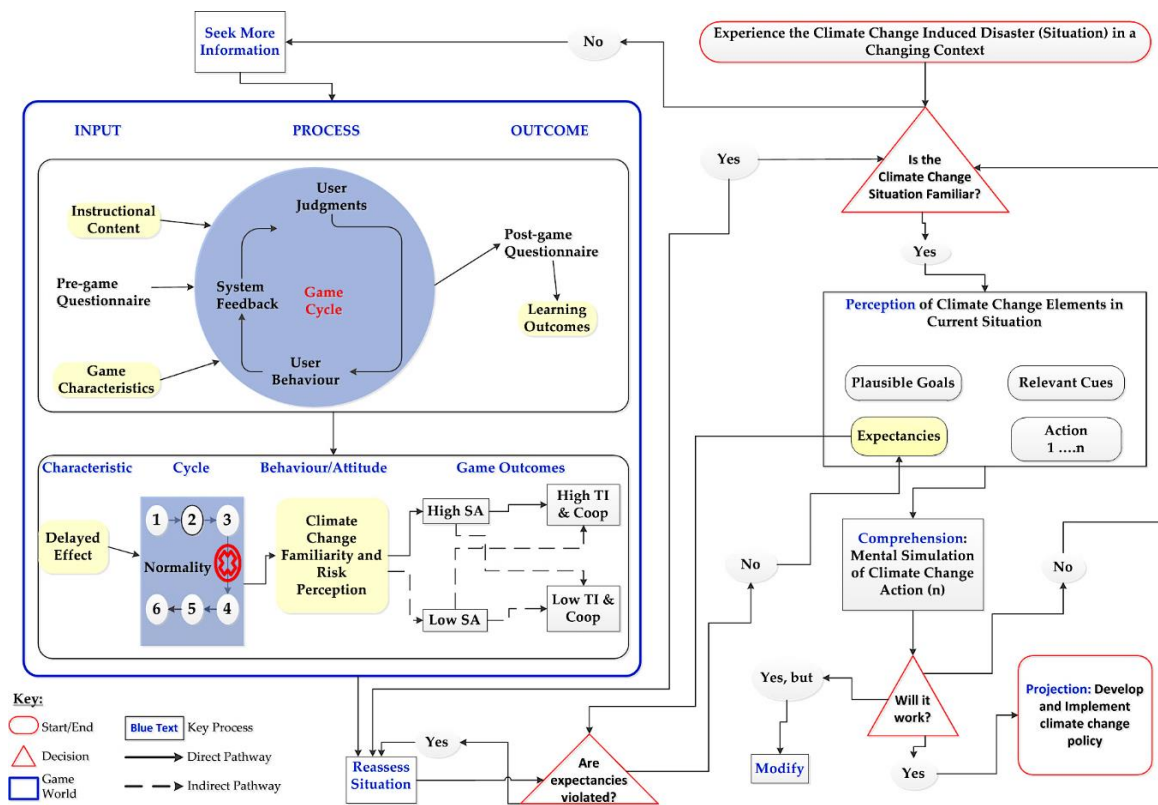
Figure 3. Nzoia WeShareIt Conceptual Framework for Situation Awareness (SA), Team Interdependence (TI), and Cooperation. Situation Awareness results are discussed in detail in Onencan [19]. Modified: Onencan [19,25].

Nzoia WeShareIt game primary focus is interdependence and cooperation [57]. We adopted the main recommendations from the Depping and Mandryk [33] namely; we incorporated competition as a positive element that enhances collaborative play and assessed cooperation and interdependence as two distinct mechanics. For cooperation, three game mechanics were incorporated, namely, shared goal, goal asymmetry, and goal synergy. For team interdependence, three game mechanics were integrated in the game design, namely, complementary role, role asymmetry, and complementary knowledge.

3.2. The Nzoia WeShareIt Game Procedures, Participants, and Measures

The gameplay consists of a series of rounds (maximum 20) and each round consists of six steps. First, players receive their resources. Second, they trade in food, wood fuel, and hydroelectric energy. Third, if applicable, players pay penalties for energy shortages. Fourth, the players invest in public services and get their (un)happy face scores. Fifth, they make water allocation decisions and buy solar panels (optional). Finally, move to the next round [23,57].

During the trading round, the players have an option to cooperate or make unilateral decisions. Unilateral decisions are possible for the resource-rich counties like Trans Nzoia and Uasin Gishu. They may choose to produce food and energy only for their county governments and not engage in any trade activities. In the trade round, players can either negotiate around the table or move around the room with their iPads looking for buyers or sellers, to meet their money, food, or energy needs. Once a trade exchange is made, the players record the transaction in their iPads [57].

The game contents include five playing fields, five information sheets, and five iPads. Most of the game activity is conducted with the help of the iPad, which is connected to the online game. The five playing fields comprise of water circles where player make water allocation decisions at the end of every round. The players can convert existing parcels into food, hydro-electric power, and nature. The game restricts allocation decisions to the number of water circles in their respective playing fields.

The players are water managers and policymakers within the Nzoia River Basin. The players are expected to collect as many smileys, as possible. The game electronically awards smileys, when a prescribed amount of food, energy, investments in public services are attained [57].

The participants were water policymakers from four local governments within the Nzoia River Basin. The four local authorities are Busia, Kakamega, Bungoma and Trans Nzoia. The participants were 35 in total (12 female and 23 male). There were seven sessions, and each of the sessions had five participants. All the participants were Kenyans from the county governments, public sector, and water companies. Participants provided informed consent before completing the questionnaires [57].

At the start of the game, participants completed the pre-game questionnaire that contained demographic questions, trust-related questions, and the consent form. Subsequently, we explained how the game process and immediately after the initiated the game. The game sessions started in the morning at 9 am, and the facilitator concludes the game at lunchtime. After the conclusion of the game, the participants fill the post-game questionnaire. Finally, we facilitated a short debriefing session, before proceeding for lunch [57].

A few questions were left blank by some participants. However, we did not find any non-compliant participants. The participants had carefully completed all the questionnaires. The final dataset contained 35 participants. The questionnaire measures for all the responses were on a 5-point-Likert (1.00 = Very Inaccurate, 2.00 = Moderately Inaccurate, 3.00 = Slightly Inaccurate, 4.00 = Moderately Accurate and 5.00 = Very Accurate). The select data, for this particular paper, did not include items that needed to be reverse coded.

3.3. Scale Statistics and Reliability Analyses

At the start of the analysis, we had 12 items for team interdependence and 11 items for cooperation. However, during the final Principal Component Analysis (PCA), three items were eliminated because they failed to load more than 0.4 in one of the components. We eliminated the three items from the overall 23 items scale, based on the PCA results. However, we retained the original component structure. Appendix A (including Table A1) explains the PCA procedure and results. We assessed internal consistency for the three scales, using Cronbach's alpha (see Table A2). After that we assessed the descriptive statistics of the team interdependence and cooperation scales. The results of the Shapiro-Wilk test and the supporting descriptive statistics are provided in Tables 2 and 3.

Table 2. Descriptive Statistics for Cooperation.

	Mean	Std. Dev
I worked with others to achieve a common interest	4.71	0.79
We managed the Nzoia water resources jointly	4.66	0.87
I enjoyed pursuing the shared goal with other counties	4.80	0.72
I am more aware of the benefits of working together	4.77	0.77
I worked with other players to be successful	4.69	0.76
I preferred to trade with our neighbors to reach our game goals	4.69	0.80
I respected other players	4.71	0.79
I got along with most of the other player	4.60	0.77
I am more aware of the need for joint action	4.69	0.76
My collaboration skills were enhanced	4.77	0.73

The means for cooperation is high confirming that the data is negatively skewed. Two hundred eighty-four of the responses perceive the contribution of the game in increasing their understanding of cooperation as "very accurate." Fifty-one preferred the "moderately accurate" category, whereas five

preferred "slightly inaccurate" category. One respondent consistently rated chose the "very inaccurate" category (ten ratings in this category in total, each for the ten dependent variables).

Table 3. Descriptive Statistics for Team Interdependence.

	Mean	Std. Dev
My knowledge on balancing food, energy, and environment increased	4.80	0.72
I felt responsible to the group for the regional strategy	4.51	0.82
I became a valuable member in the game	4.63	0.77
I established a role that increased interactions with others.	4.80	0.76
We made water allocation decisions together	4.51	0.89
We gave feedback on each other's contribution	4.46	0.85
I learned the importance of sharing information	4.77	0.73
I interacted with others to achieve game goals.	4.60	0.91
I respected other players	4.60	0.88
I developed social connections	4.54	0.85

The means for team interdependence is lower than cooperation. The team interdependence dataset is also negatively skewed. Two hundred sixty-three of the responses perceive the contribution of the game in increasing their understanding of team interdependence as "very accurate." Sixty-three preferred the "moderately accurate" category, whereas 13 and 1, preferred "slightly inaccurate" "moderately inaccurate" categories, respectively. The same respondent consistently chose "very inaccurate" category for cooperation, provided similar results for team interdependence.

After assessing the descriptive statistics, we conducted a graphical test, using boxplot chart histograms to identify significant outliers (Figure 4).
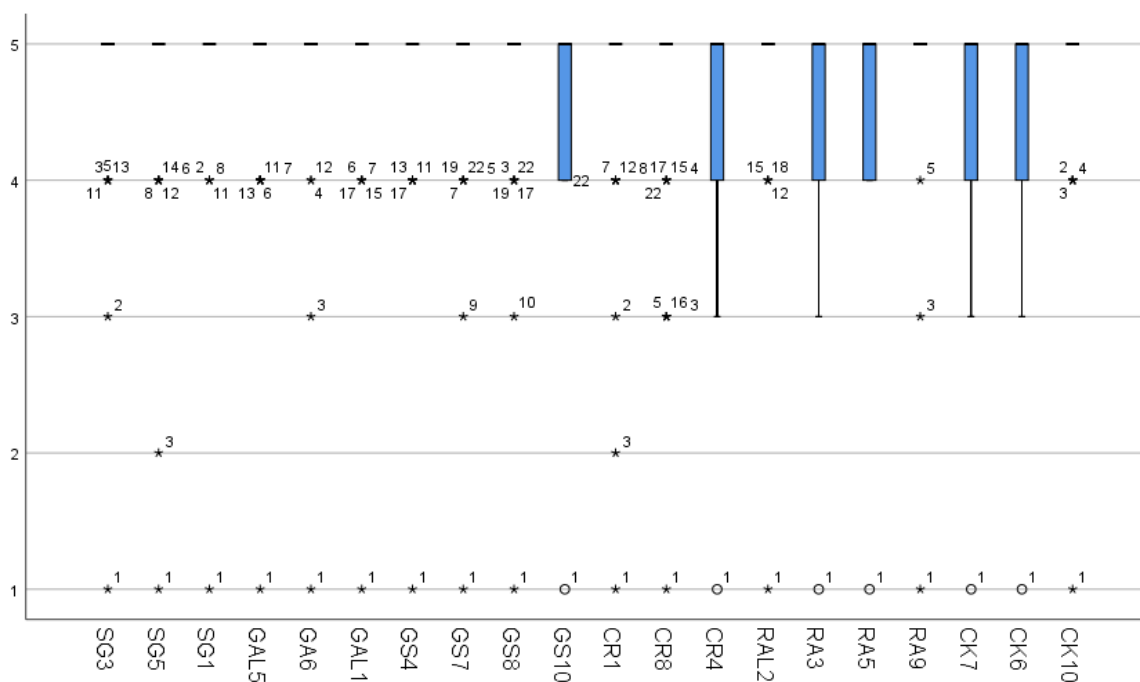

Figure 4. Boxplot representation of the outliers or extreme values for the team interdependence and cooperation scale, distributed by the 20 dependent variables.

The presence or absence of outliers informed the choice of method for the statistical analyses. We identified multiple outliers. A number of these outliers appear to be extreme because of their distance from the mean. Appendix B provides the details concerning the assumption testing and the analyses we made regarding the significant outliers.

3.4. Data Analyses

We performed data analyses using the Statistical Package for the Social Sciences (SPSS) version 25 and the statistical package for Excel (XLSTAT) version 19.7. Overall, a few questions were left blank by some participants. However, we did not find any non-compliant participants. The final dataset contained 35 participants. Having decided to maintain the outliers, we still had a problem; these extreme values may inflate the within-group variability when we conduct parametric results, thus affecting the significance assessment. Also, the data series is negatively skewed with most of the respondent ratings between 3 and 5 and very few ratings in the 1 and 2 range. Since we decided to keep the outliers, we selected the Chi-Square test for goodness-of-fit, which is useful for Likert scale data. The test is useful in comparing the observed distribution (sample distribution) with a theoretical distribution. Also, we conducted a separate one-way between-subjects ANOVA with Friedman's nonparametric test on the same data.

PCA data were performed using XLSTAT. We used PCA to visualize correlations amongst our original variables and between these variables and the components, to improve the questions asked in the cooperation and team interdependence sub-scale, to inform future research.

4. Results

This section focuses on the results of the Chi-Square test for goodness-of-fit and the one-way between-subjects ANOVA with Friedman's nonparametric tests. The Chi-Square test for goodness-of-fit null hypothesis is: there is no significant difference between the observed distribution for the cooperation and team interdependence dependent variables and the theoretical distribution (H0). The alternative hypothesis for the Chi-Square test for goodness-of-fit is: there is a significant difference between the observed distribution for the cooperation and team interdependence dependent variables and the theoretical distribution (H1). The theoretical distribution corresponds to a situation of indifference where the responses are at mid-point (3) in the 5-point Likert Scale. We reject the null hypothesis if asymptotic significance < 0.05. The *p*-value for all the ten variables is lower than 5% (asymptotic significance < 0.05).

4.1. Chi-Square Test for Goodness-of-Fit

We reject the null hypothesis for the cooperation results. There are significant differences between the observed frequencies and the expected frequencies (Table 4).

Table 4. Chi-Square Test for Goodness-of-Fit Results for Cooperation.

	SG3	SG5	SG1	GS4	GS7	GS8	GS10	GAL5	GAL1	GA6
Chi-Square	63.17 [a]	57.69 [a]	48.23 [b]	31.77 [b]	57.69 [a]	63.17 [a]	23.03 [b]	31.77 [b]	43.60 [b]	75.51 [a]
df	3.00	3.00	2.00	2.00	3.00	3.00	2.00	2.00	2.00	3.00
Asymp. Sig.	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005

[a] 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 8.8. [b] 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 11.7.

The *p*-value for all the ten team interdependence variables is lower than 5% (asymptotic significance < 0.05). Therefore, we reject the null hypothesis. There are significant differences between the observed frequencies and the expected frequencies (Table 5).

Table 5. Chi-Square Test for Goodness-of-Fit Results for Team Interdependence.

	RAL2	RA3	RA5	RA9	CK7	CK6	CK10	CR1	CR8	CR4
Chi-Square	48.23 [a]	34.37 [b]	25.60 [a]	82.27 [b]	37.57 [b]	29.80 [b]	43.60 [a]	73.14 [c]	51.29 [b]	38.71 [b]
df	2.00	3.00	2.00	3.00	3.00	3.00	2.00	4.00	3.00	3.00
Asymp. Sig.	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005

[a] 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 11.7. [b] 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 8.8. [c] 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 7.0.

4.2. One-Way Between-Subjects ANOVA with Friedman's Nonparametric Test

We employed a one-way between-subjects ANOVA with Friedman's nonparametric test on the independent ratings on the effect of team interdependence mechanics, independent ratings on the effect of the cooperation mechanics and the combined ratings of all the game effects. This revealed a significant increase in understanding of team interdependence: X^2 (34, N = 12) = 32.26, p = 0.001. It also revealed a significant increase in understanding of cooperation: X^2 (34, N = 11) = 23.42, p = 0.009.

5. Discussions

In the discussion section, first, we analyze the findings of recent articles on the SDG 6.5.2 indicator. Then, we explain the contribution and new results of the current study. Finally, we make policy recommendations to WASREB based on the literature review and the findings of the present study.

5.1. The Contribution of Recent Studies on SDG 6.5.2 Indicator

By reducing complexity, UN-Water chose the path of selecting targets that are quantifiable [37,40]. The advantage of this approach is timely and cost-effective measurement of performance [29]. Unfortunately, quantifiable targets alone cannot measure the complexity facing most shared river basins [37,67]. Hussein state that "quantitative methodologies are generally not able to capture nuances and different shades, forcing towards fixed and set categorizations" [37] (p. 5). Before any written agreement, most countries implement hybrid forms of cooperation, over extended periods [68,69]. Therefore, legal agreements is a narrow definition of cooperation [69]. SDG 6.5.2 considers agreements and arrangements as a precondition to cooperation [29], thus narrowing the scope of transboundary water cooperation and affecting the sustainability of joint basin actions [37]. Table 6 summarizes the identified policy gaps of SDG indicator 6.5.2, the proposed action(s) and the authors who contributed to the studies.

Table 6. Relevant articles contributing to the study the SDG 6.5.2 Policy Gaps.

Indicator 6.5.2 Gap	Policy Gap	Proposed Action (s)	Relevant Article
1. Team Interdependence	• Interdependence of states that share a basin and team interdependence of transboundary groups not measured	• Measure team interdependence separate from cooperation	Depping and Mandryk [33]
2. Ambiguous, Diverse Interpretations & Results	• Criteria to measure operationality of cooperation is ambiguous and confusing • Different countries interpret meetings and the joint river basin management differently leading to non-comparable results.	• Revise ambiguous indicators. • Develop standard and harmonized indicators and improve monitoring to increase comparability of data	Hussein, Menga and Greco [37] de Chaisemartin [60], McCracken and Meyer [40], (p. 9) Ortigara, et al. [70], (p. 8)
3. Operational Cooperation & Pre-Cooperation Phase	• Limited to official/formal cooperation thus leaving out the pre-cooperation phase. • Operational cooperation requires that a riparian state meets all the four criteria, to qualify. • The binary indicator fails to take account of progress and stepwise cooperative processes	• Incorporate formal, informal and technical state deliberations (preoperational phase) • Develop indicators that are not binary and take account of stepwise cooperative processes.	Hussein, Menga and Greco [37] McCracken and Meyer [40], (pp. 9–10) de Chaisemartin [60], (p. 20) McCracken [71] Tarlock [72]

Table 6. *Cont.*

Indicator 6.5.2 Gap	Policy Gap	Proposed Action (s)	Relevant Article
4. Sub-basin Operational Coop Complexity	• Exclusion of cooperation actions and arrangements that do not cover the entire basin. • Sub-basin results of SDG 6.5.2 are not comparable. • Failure to assess substantive issues, which make data amalgamation inaccurate and unreliable.	• Include a provision to address cooperative arrangements not covering the whole basin. • Develop indicators that measure substantive elements of cooperation.	McCracken and Meyer [40], (p. 10). Hussein, Menga and Greco [37]
5. Cooperation Quality is not measured	• Exclusion of process and input indicators, thus not measuring the quality of cooperation. • Formal agreements cannot measure ongoing hydropolitics. • Outcome indicators fail to capture the progressive improvements made.	• Use of indices analyzed and suggested [40,60] • A scale of Cooperation proposed [71,72] • Qualitative analyses including hydro-political assessments and discourse analyses proposed [37].	Hussein, Menga and Greco [37] McCracken and Meyer [40] de Chaisemartin [60] Ortigara, Kay and Uhlenbrook [70] McCracken [71] Tarlock [72]
6. Hybrid Cooperation & Non-State Discourses	• Informal cooperative arrangements cannot be measured under indicator 6.5.2 [37]. • Cooperative actions by politicians (political statements), media, NGOs, civil societies, and development partners, excluded.	• Include qualitative indicators to assess state and non-state discourses and hybrid cooperation.	Hussein, Menga and Greco [37]
7. Power Dynamics in a Transboundary Basin	• Fails to take account of the power dynamics that exist in some river basins.	• Develop indicators that unveil concealed power asymmetries and domination.	Hussein, Menga and Greco [37]
8. Distinguish between Good & Bad Cooperation	• Indicator 6.5.2 does not make a distinction between good and bad cooperation.	• Distinguish good from bad cooperation and measure the quality of cooperation.	Hussein, Menga and Greco [37]
9. Overlaps, Inequity, and Inequality Issues	• Area as the unit of analysis distorts the overall picture due to many overlaps (basin & sub-basin levels, surface water & transboundary aquifers) • A single jacket approach to address existing complexities has led to inequitable outcomes	• A shift from the area as the unit of analysis towards use of other units of measurement • Water volume/# of water users/# of agreements proposed [40], (p. 10) • Area of the riparian state in the basin/volume of water proposed [60]	McCracken and Meyer [40] de Chaisemartin [60]
10. Data Quality, Reliability & Availability	• Binary (yes or no) nature of the indicator makes it difficult to grasp the actual cooperation situation. • There is no robust mechanism to check the quality of the data which affects the reliability of the data. • Ambiguous indicators lead to multiple interpretations and data that is not comparable nor interoperable. • Lack of disaggregation of the data makes it difficult to assess the actual situation • Basins and aquifers reported spatial data collection problems.	• Revise the binary indicators, address the ambiguous issues and disaggregate the data [40]. • Develop standard and harmonized indicators and improve monitoring to increase the comparability of data [70], (p. 8) • Address data availability and financial related data collection challenges [70].	McCracken and Meyer [40], (p. 9). Ortigara, Kay and Uhlenbrook [70], (p. 10) McCracken [71]

Based on Table 6, there is a consensus regarding the identified gaps to monitoring water cooperation using SDG Indicator 6.5.2. The difference is on the proposed solution for gap number five (how to measure the quality of cooperation). Most of the studies recommend measuring cooperation with indices [40,60,70–72], while one research suggest qualitative data [37]. Examples of indices/scales that have been proposed by various researchers are:

1. A scale of Cooperation presented in the GWP TEC Background paper by McCracken [71];
2. A scale of Cooperation presented in the GWP TEC Background paper by Tarlock [72];
3. Water Cooperation Quotient by the Strategic Foresight Group, in Water Cooperation [73];

4. Transboundary Freshwater Dispute Database (TFDD) Basin at Risk (BAR)/International Water Intensity Scale presented by the Oregon State University (OSU), in the Transboundary Freshwater Dispute Database [74]; and

5. Transboundary Water Interaction Nexus (TWINS)] by Zeitoun and Mirumachi [55].

We compared these scales and the Benefit Sharing Model [41,75,76], and concluded that most of the scales are robust in assessing benefits to the river and because of the river. Only the water cooperation scale addresses the benefits from the river, and none of the scales is adequate to evaluate benefits beyond the river.

According to Sadoff and Grey [76], (pp. 393–395), benefits to the river creates opportunities for improved river flow, water quality, conservation of soil, protection of biodiversity, and ensuring sustainability. Increased benefits to the river also reduce pollution, land degradation, loss of biodiversity, drying of wetlands and unmanaged watersheds. Benefits from the river, lead to increased food production, hydropower production, and sharing the energy through interconnection lines, increased river navigation, increased recreation and tourism, and improved flood-drought management [76], (pp. 395–397). Improved cooperation may facilitate the process of reducing costs because of the river [76], (pp. 398–399). Most of the river basins barely engage in cooperation until a conflict arises [1,23,67]. Therefore, benefits because the river helps the riparian states to agree on a governance framework that shifts unilateral actions towards joint management and development [68]. Benefits because the river also reduce the food and energy production costs, through an agreement on joint cost and benefit sharing [49,69]. Finally, the process of developing a governance framework to support the regional integration of the riparian nation-states increases the benefits derived beyond the river [76], (pp. 399–400).

Hussein, Menga, and Greco [37] propose qualitative analyses that incorporate discourse analyses and hydro-political studies. The advantage of qualitative analyses and indices is the ability to track progress, explicitly. McCracken [71], (p. 66) explain that indices and qualitative data are beneficial because they provide more flexibility in cooperative actions, assess political will, establish the socio-economic context within which cooperation occurs, support step-wise cooperative processes, and acknowledge the cooperative actions of non-state actors. However, McCracken [71], (p. 65) highlight that qualitative data "does not present a single value for each country, which does not meet the needs for global SDG monitoring." To address this gap, we propose collection of both qualitative and quantitative data. In addition, a stronger SDG monitoring mechanism should be established. Therefore, qualitative data will be used to substantiate and justify the quantitative data.

5.2. Contribution of the Nzoia WeShareIt Water Policy Game

The Nzoia WeShareIt game results indicate that a water policy game is a useful experiential learning tool on team interdependence and cooperation. Both the Chi-Square test for goodness-of-fit and the one-way between-subjects ANOVA with Friedman's nonparametric tests confirm that the game positively contributed to the planned learning outcomes on team interdependence and cooperation. The two sub-scales indicators are a subjective assessment of respondents' game learning outcomes. These game-specific questions need not to be adopted by WASREB in its normal operations. The water policy game's contribution is testing and suggesting an innovative WSP capacity development on team interdependence and cross-county cooperation. WASREB could use the results to develop similar social innovations and assess learning outcomes.

Of primary importance are the cooperation and the team interdependence game mechanics that we used to design the game. These mechanics may form the foundation for the design of Kenyan specific indicators to measure cooperation and team interdependence. For cooperation, we propose the introduction of shared goal, goal asymmetry, and goal synergy indicators. For team interdependence, we propose the introduction of complementary knowledge, role asymmetry, and complementary roles indicators. The practical application of the six mechanics may entail:

1. Shared Goal indicator(s) for measuring cross-county cooperation. An indicator to measure whether county governments within a given basin have a shared goal to sustainably manage the water resources for the benefit of all the basin residents.

2. Goal asymmetry indicator(s) for measuring cross-county cooperation. Each county government has different water availability levels and demand. Therefore, the current KPI indicators should be revised to avoid inequitable outcomes. WASREB should integrate the volume of water, water demand and supply indicators into the current KPIs. Besides, WASREB should encourage, support, finance and measure cross-county water partnerships.

3. Goal synergy indicator(s) for measuring cross-county cooperation: WASREB should measure the interactions between WSPs that promote interdependence and cooperation. In Nzoia Basin, there is one cross-county WSP, NZOWASCO [2] (p. 74). WASREB should develop cross-county collaborations indicators directly linked to SDG 6.5.2 monitoring financing advanced to WSPs.

4. Complementary knowledge indicator(s) for measuring cross-county team interdependence: To enhance team interdependence WASREB may develop joint reporting mechanisms where WSPs need information from other county governments to comply with the KPIs. For instance, indicators that measure the impact of WSP actions on other riparian county governments. This indicator will require the various WSPs within the shared drainage basin to agree on the information submitted to WASREB to avoid conflicting information. Through their interactions and reporting, cross-county team interdependence will be enhanced.

5. Role asymmetry indicator(s) for measuring cross-county team interdependence: WASREB should develop indicators that take account of the different roles that county governments from various geographical zones have to undertake. The county governments in urban centers where residents solely rely on WSP for water and sanitation services have different roles from rural-based county governments. Therefore they may not require the same number of staff. Some WSP have no water distribution role during prolonged dry seasons because they have no water to distribute. All these complexities should be included in the revised WASREB indicators.

6. Complementary roles indicator(s) for measuring cross-county team interdependence: Different county governments have different comparative advantages. Therefore, WASREB KPIs should acknowledge the complexity and revise their current KPIs to reflect this. For instance, the cross-county collaboration between Bungoma and Trans Nzoia that led to the formation of NZOWASCO, creates complementary roles. Trans Nzoia is an upstream riparian government. Therefore their role is to extract water resources from Mt. Elgon and Cherangany Hills, treat and distribute downstream. The role of the Bungoma county government is to store the water resources and distribute to its residents. These complementary roles need to be acknowledged and assessed differently in the revised WASREB KPIs.

5.3. From Policy to Practice: Recommendations to WASREB

WASREB monitors the implementation of SDG 6 at the WSP level. However, to foster cross-county cooperation, we recommend monitoring SDG 6 at the basin level. WASREB should encourage and support cross-county cooperation and continuosly monitor cross-county cooperative actions. Furthermore, if WASREB decides to use the UN-Water SDG 6.5.2 indicator's methodology to assess cross-county cooperation, there is a need for the following adjustments:

1. Disaggregate the SDG 6.5.2 indicators and have a clear definition of terms.
2. Develop qualitative and process-based indicators for operational cooperation.
3. Develop clear indicators to measure team interdependence and cooperation as two distinct indicators.
4. Develop indicators that take account of the basin complexities and peculiarities.
5. Use a mixed methods approach to measure cooperation, that incorporates substantive elements of the formal agreements to support the current quantitative form of data collection. An SDG

6.5.2 index should be developed that complies with the cooperation continuum as described in Sadoff and Grey [75,76]

6. Develop indicators to measure hybrid forms of cooperation, and expand the measurement of Indicator 6.5.2 to non-state actors. WASREB should use discourse analysis methods to measure the emerging discourses and how they impact on transboundary water cooperation

7. Use mixed methods to assess power dynamics within a given basin. Since the power dynamics change, the results should be regularly updated.

8. Institute a peer-review mechanism to distinguish good from bad cooperation. Also, the qualitative data and the study of power dynamics would provide valuable information on whether the finalized agreement is masking bad cooperation or cooperation is good.

9. The unit of analysis should be carefully selected to ensure that it does not propagate inequities and deepen complexity. Area as a unit of analysis is misleading and should be replaced.

10. Institute mechanisms to promote data sharing, check on the reliability of the data being provided. Also, non-availability of data needs to be urgently addressed. Disaggregate data to identify whether WSPs are extracting water from surface water bodies or aquifers.

11. Before licensing WSPs, WASREB should require that they conduct detailed analyses of surface and groundwater resources. Thereafter, WASREB should use the data to develop an integrated information system, jointly managed by all WSPs. To encourage data sharing WASREB should develop indicators to measure the quality of data shared and contributions of a WSP to the overall information management system.

6. Concluding Remarks

The paper conducts a detailed literature review of SDG indicator 6.5.2 that measures transboundary water cooperation. The focus of the paper is to support local monitoring of water cooperation, especially at the cross-county governance level. The target institution is the Water Services Regulatory Board (WASREB), which monitors the implementation of SDG 6 at the local level. Based on the analyses of previous studies and the results of the Nzoia WeShareIt game, the paper makes some recommendations that WASREB may incorporate into their current KPIs, to enhance cross-county cooperation.

To improve indicator 6.5.2, this paper introduces a mixed method approach and a set of indicators that focus on the quality of cooperation. The process and content are given more weight than formal documents that may or may not be produced to cement the relationship. The proposed methodology moves away from the overall assessment of outputs towards a subjective assessment of the quality of cooperation, by both state and non-state actors. The mixed method approach lacks a single value of measurement, thus not meeting the requirements of a global indicator. However, it is important that the mixed method approach is adopted and the single value indicators can be later extracted for the global monitoring of SDG 6.5.2. This approach ensures that the data submitted for global monitoring of SDG 6.5.2 is substantiated by high-quality data that is both quantitative and qualitative.

Also, we introduce a critical aspect of cooperation—team interdependence—to enrich cooperative relations and check power asymmetries. Team interdependence indicators measure the extent of connectivity between the different riparian local governments. Team interdependence is enhanced when there is a valued shared goal and the cross-county communications and actions are targeted towards achieving the shared goal.

Another contribution of the study is an assessment of the value addition of a water policy game in enhancing cross-county cooperation. Nzoia WeShareIt game results indicate that gaming is a promising method for encouraging cross-county cooperation in a region prone to unilateral actions by the local authorities. The game created a learning space where the policymakers could test various water cooperation options and identify the configurations that would create benefits for all the riparian county governments.

This study was the first application of SDG 6.5.2 in the Kenyan county government level. There is a need for further assessments aimed at deepening understanding on existing global and local indicators and their impact on sustainability and equity. Since the application was confined to a water policy game, there is a need for actual real-life application of the recommendations to assess whether the gaming environment was an accurate reflection of the reality in Nzoia river basin. Future studies should also focus on how the findings can be replicated to larger basins, including Lake Victoria and the Nile Basin.

Author Contributions: A.M.O. Conceptualized the article, designed the Survey Monkey questionnaire, conducted the game sessions, collected the data, undertook the in-game performance measurements analysis and the SPSS analyses in the context of team interdepedence and trust, wrote the original draft and was actively involved in the draft preparation, content visualization, draft improvement and the incorporation of comments from the second and third authors and the reviewers. B.E. was actively involved in the policy game design and the game-testing sessions, improved the initial conceptualization, methodology, actively participated in the validation process, and was also responsible review & editing, and formal analysis. B.V.d.W. was actively involved in the policy game design, mobilized the financial and technical resources to design and implement the policy game, actively participated in the validation process, and was also responsible for review & editing, supervision, project administration, and formal analysis.

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

Appendix A. Principal Component Analysis (PCA) Analyses and Results

Three principal component analyses (PCA) were undertaken to identify the variables that load heavily on the principal components.

Appendix A.1. Data Screening and Summarised Statistics

For all the assessments, the PCA results excluded three participants for missing data, ID 12, 15, and 29. The data for component analysis contained a final sample size of 32 (using listwise deletion), with 30 variables.

The first result is the summarised statistics for each variable. The mean results are very high for all the datasets indicating that the participants rated the game results very positively. For cooperation, the lowest mean is 4.4 for the scale item C_2 and the highest is 4.8 for the scale item C_1. The standard deviation for this scale ranges from 0.72 for C_1 to 0.87 for C_5. For team interdependence, the lowest mean is 4.4 for the scale item I_2 and the highest is 4.8 for the scale item I_9 and L2I. The standard deviation for this scale ranges from 0.72 for L2I to 0.91 for I-1. For overall learning results, the lowest mean is 4.5 for the scale item L_7, I_3, I_6 and I_7 and the highest is 4.8 for the scale item L_2, I_9 and C_1. The standard deviation for this scale ranges from 0.72 for L_2 and L_8 to 0.91 for I-1.

Initially, the factorability of the 11 cooperation, 12 team-interdependence, and 30 overall learning items, was examined. The PCA analyses used well-recognized criteria to assess the factorability of a correlation. Firstly, all the 30 items correlated at least 0.47 with at least one other item, suggesting reasonable factorability. We observed high inter-item correlations with ranges between 0.6 and 0.93. The cooperation correlation matrix indicates that only three correlations are in the 0.5 range and one in the 0.4 range. The team-interdependence correlation matrix has all the inter-item correlations above 0.6 except four correlation values related to I_2, which are below 0.6 and above 0.49.

The three scales passed the Kaiser-Meyer-Olkin measure of sampling adequacy. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.83 for cooperation, 0.87 for team-interdependence and 0.78 for the overall learning. The three results are above the recommended value of 0.6.

The cooperation and team interdependence scales indicate significant results for Bartlett's test of sphericity. The results are as follows: cooperation (χ^2 (55) = 449.01, $p < 0.0001$) and team interdependence (χ^2 (66) = 496.01, $p < 0.0001$). The diagonals of the anti-image correlation matrix were all over 0.5, supporting the inclusion of each item in the PCA. We could not compute Bartlett's sphericity test on the learning scale because of multicollinearity between the selected variables.

The communalities were all above 0.3, further confirming that each item shared some common variance with other items. Given the positive results in all the overall indicators, we proceeded to conduct a PCA, with all the 30 items.

Appendix A.2. The Underlying Structure Model of Team Interdependence, Cooperation, and General Learning Outcomes

PCA was used to identify the principal components underlying the team interdependence, cooperation, and overall learning assessment scales. PCA was conducted using the XLSTAT software, and the selected PCA type was "Correlation." The maximum number of filter factors were five. We based standardization on the sample size (n) of 35 participants. The select method of rotation was Varimax (Kaiser normalization), and the number of components rotated was two. The biplot type was correlation biplot with an automatic coefficient. We also conducted bootstrap observations with a sample size of 50.

The initial eigenvalues for the three scales favored a two-component solution. For cooperation, the initial components were 11 in total. Initial eigenvalues indicate that the first component explains 74.98% of the variance, the second component 5.73% of the variance, and a third component 5.03% of the variance. The other eight components had eigenvalue variability (%) ranging from 4.43% to 0.17%, and all the eight components explaining 14.26%.

For team-interdependence, the initial components were 12 in total. Initial eigenvalues show that the first component explains 75.94% of the variance, the second component 6.28% of the variance, and a third component 4.21% of the variance. The other nine components had eigenvalue variability (%) ranging from 3.79% to 0.20%, and all the nine components explaining 13.57%.

For overall learning, the initial components were 30 in total. The initial eigenvalues showed that the first component explained 72.16% of the variance, the second component 3.96% of the variance, and a third component 3.6% of the variance. The other 27 components had eigenvalue variability (%) ranging from 3.3% to 0.001%, and all the 27 components explaining 20.29%.

We preferred a two-component solution which explained 80.71%, 82.23%, and 76.12% of the variance for the cooperation, team- interdependence and overall scale, respectively, based on three reasons. First, we noticed the leveling off of eigenvalues on the scree plot after the first component. Second, there was an insufficient number of primary loadings after the first component. Third, we experienced difficulty in interpreting the second component and following components.

Appendix A.3. Analysis of the PCA Scree Plots and Biplots

The scree plots indicate that the first component explains the most significant portion of the eigenvalues.

- The scree plots also visualize a steep drop and leveling off to form a straight line from after the first component.
- The scree plot helped in the selection of the two-component solution for all the three scales. We visualize this two-component solution which explained 80.71%, 82.23%, and 76.12% of the variance, in the form of three bi-plots. Since the first component explains an adequate amount of variation, then we used the first two principal components to select the relevant bi-plot for each scale, that would proceed for further analysis.

The bi-plots shows the following:

- The data points are spread out indicating that the responses are reliable.
- All the scale items have large positive loadings on component 1.
- There are significantly positive and negative loadings on component 2.
- Component 1 explains the highest variability, with component 2 being very small due to the significant negative loadings.

- "ID8" data point in the upper-left hand corner and "ID24" lower left-hand corner may be outliers.
- Scale items "C_2" and "I_2", are outliers and do not contribute to a simple component structure.

The bi-plots after the Varimax rotation (Kaiser normalization) indicate the following:

- All the scale items have large positive loadings on both component 1 and 2.
- Component 1 and 2 explain almost the same amount of variability, with no negative loadings.
- "ID8" data point in the lower left-hand corner and "ID24" lower right-hand corner may be outliers.
- Finally, scale items "C_2" and "I_2", are outliers and do not contribute to a simple component structure.

We filtered the components to a maximum number of five. After that, we conducted a rotation using the varimax (Kaiser normalization) rotation of the component loading matrix.

Appendix A.4. Improvement of Existing Scale

All the 30 items met the minimum criteria of having a primary component loading of 0.4 or above and cross-loading of 0.3 or above. However, during the final PCA steps, a total of five items were eliminated because they did not contribute to a simple component structure and failed to load more than 0.4 in one of the components. The item "L_9" did not load above 0.33 on D2. The item "L_10" did not load above 0.32 on D2. The item "I_2" did not load above 0.27 on D1. The item "C_2" did not load above 0.4 and 0.5 on D1 and D2 respectively. The item "C_9" did not load above 0.31 on D1. Also, "C_2" and "I_2", were outliers and did not contribute to a simple component structure.

We eliminated five items from the overall thirty items scale, based on the PCA results. However, we retained the original component structure. Overall, these analyses indicated that two distinct components were underlying the game assessment tool for the revised version of the Nzoia WeShareIt game assessment tool for team interdependence, cooperation, and general learning.

A PCA of the remaining 25 items, using varimax (Kaiser normalization) rotation of the component loading matrix, was conducted, which explained 83.58%, 83.03%, and 79.02% of the variance for the cooperation, team-interdependence, and overall scale, respectively. A two-factor structure for 25 out of the 30 items was evident, based on a principal components exploratory factor analysis, using a rotation using the varimax (Kaiser normalization) rotation of the component loading matrix.

Appendix A.5. Assessment of the Scales Internal Consistency

The reliability analyses indicate optimal internal consistency. All the components were internally consistent with extremely high alphas. All items of the revised 25 item scale had primary loadings over 0.5, and only two items had a cross-loading below 0.4. However, these items had a strong primary loading of 0.83 for both items "L2I" and "L_6". Table A1 presents the component loading matrix for this final solution.

Table A1. Component Loadings based on PCA with Varimax rotation for 25 items with details of the sub-scale indicators/question contained in the post-game questionnaire, for each item.

Item		Sub-Indicators for Cooperation, Interdependence and General Learning	D1	D2	Communality
Cooperation	C_1	Shared Goal: I enjoyed pursuing the shared goal with other counties	0.75	0.54	0.84
	C_3	Shared Goal: I worked with others to achieve a common interest	0.88	0.40	0.77
	C_4	Goal Synergy: I worked with others to be successful	0.60	0.69	0.83
	C_5	Shared Goal: I managed the Nzoia water resources jointly	0.73	0.44	0.69
	C_6	Goal Asymmetry: I am more aware of the benefits of working together	0.75	0.53	0.82
	C_7	Goal Synergy: I valued cooperation over competition	0.53	0.66	0.71
	C_8	Goal Synergy: I respected others	0.58	0.66	0.76
	C_10	Goal Synergy: I got along with most of the others	0.85	0.43	0.78
	L1C	Goal Asymmetry: I am more aware of the need for joint action	0.40	0.88	0.77
	L5C	Goal Asymmetry: My collaboration skills were enhanced	0.41	0.86	0.84
Interdependence	1_1	Complementary Role: I interacted with others to achieve basin goals.	0.41	0.84	0.76
	1_3	Role Asymmetry: I felt responsible to the group for the regional strategy	0.76	0.50	0.80
	I_4	Complementary Role: I developed social connections	0.49	0.78	0.78
	I_5	Role Asymmetry: I became a valuable member in the basin	0.80	0.42	0.77
	I_6	Complementary Knowledge: I gave feedback on each other's contribution	0.44	0.84	0.79
	I_7	Complementary Knowledge: I shared information and encouraged joint action	0.64	0.58	0.74
	I_8	Complementary Role: I discussed progress with others after every meeting	0.72	0.48	0.73
	I_9	Role Asymmetry: I established a role that increased interactions with others	0.75	0.57	0.88
	I_10	Complementary Knowledge: I learnt the importance of sharing information	0.71	0.61	0.87
	L2I	Role Asymmetry: I increased their knowledge of balancing food, energy, and environment.	0.83	0.34	0.71
	L_1	I am more aware of the need for joint action	0.72	0.52	0.79
	L_2	I increased my knowledge on balancing between food, energy, and environment	0.77	0.44	0.79
	L_3	My trust of others has increased	0.72	0.46	0.73
	L_4	I enhanced my decision-making skills	0.60	0.75	0.92
	L_5	I enhanced my collaboration skills	0.58	0.73	0.86
	L_6	I enhanced my conflict-management skills	0.28	0.83	0.76
	L_7	I am more aware of the impact of climate-change induced disasters	0.61	0.49	0.62
	L_8	I am more aware of the need for trust-building	0.82	0.39	0.82

Table A1. *Cont.*

Item		Sub-Indicators for Cooperation, Interdependence and General Learning	D1	D2	Communality
	C_1	Shared Goal: I enjoyed pursuing the shared goal with other counties	0.79	0.45	0.83
	C_4	Goal Synergy: I worked with others to be successful	0.84	0.37	0.85
	C_5	Shared Goal: I managed the Nzoia water resources jointly	0.65	0.51	0.68
	C_6	Goal Asymmetry: the I more aware of the benefits of working together	0.72	0.55	0.82
	C_7	Goal Synergy: I valued cooperation over competition	0.59	0.59	0.70
	C_8	Goal Synergy: I respected another s	0.82	0.38	0.81
	C_10	Goal Synergy: I got along with most of the other	0.80	0.39	0.80
Learning, Cooperation, and Team Interdependence	I_1	Complementary Knowledge: I interacted with others to achieve basin goals.	0.43	0.78	0.80
	I_3	Role Asymmetry: I felt responsible to the group for the regional strategy	0.53	0.70	0.77
	I_4	Complementary Role: I developed social connections	0.38	0.84	0.85
	I_5	Role Asymmetry: I became a valuable member in the game	0.69	0.53	0.77
	I_6	Complementary Knowledge: I gave feedback on each other's contribution	0.43	0.79	0.80
	I_7	Complementary Knowledge: I shared information and encouraged joint action	0.57	0.64	0.74
	I_8	Complementary Role: I discussed progress with others after every round	0.48	0.68	0.69
	I_9	Role Asymmetry: I established a role that increased interactions with others	0.66	0.67	0.89
	I_10	Complementary Knowledge: I learnt the importance of sharing information	0.65	0.68	0.88

Appendix A.6. Revision of Component Loadings

Most of the primary loadings were high, indicating strong factors. However, there is a possibility of strengthening the current loadings for the two components. Once we identified the items with low primary loadings, we made revisions of the scale items.

Currently, there are three items with primary loading that are lower than 0.7. The first item is "I_7" with a primary loading of 0.64. The scale item reads: "we shared information and encouraged joint action." The two terms "shared information" and "joint action" is too broad and may not be easy to relate. Also, the item should only measure one variable, not two. This item could be revised to read: "we made water allocation decisions together."

The second item is "L_7" with a primary loading of 0.61. The scale item reads: "I am more aware of the impact of climate-change induced disasters." The two terms "impact" and "climate-change induced disasters" are complex and broad and may not be easy to relate to. This item could be revised to read: "I am more aware of the consequences of the drought on the Nzoia basin."

The third item is "C_7" with a primary loading of 0.59. The scale item reads: "I valued cooperation over competition." The two terms "cooperation" and "competition" are too broad and may not be easy to relate to. This item could be revised to read: "we preferred to trade with our neighbors to reach our game goals".

Further research, beyond the scope of this paper, may entail assessing whether to adding more scale items to replace the deleted scale items, would improve scale performance.

Appendix B. Additional Statistical Analyses Information

Appendix B.1. Adjustments Made after PCA

The alphas were optimal at 0.97 for Team Interdependence (10 items), and 0.97 for cooperation (10 items). The results of the Cronbach's Alpha if Item Deleted indicated that we could not achieve considerable increases in alpha for any of the three scales, by eliminating any of the remaining items (Table A2). The reliability analyses indicate optimal internal consistency. All the components were internally consistent with extremely high alphas.

Table A2. Previous and Revised Reliability & Scale Statistics for the Select Assessment Scales.

Purpose	Reliability & Scale Statistics	# of Items
Team Interdependence	Previous: Cronbach's α = 0.97, M = 55.31, SD = 8.55 (N = 35, 0 excluded)	12
	Revised: Cronbach's α = 0.97, M = 46.23, SD = 7.23 (N = 35, 0 excluded)	10
Cooperation	Cronbach's α = 0.97, M = 51.49, SD = 7.36 (N = 35, 0 excluded)	11
	Revised: Cronbach's α = 0.97, M = 47.79, SD = 6.84 (N = 35, 0 excluded)	10

Appendix B.2. Test of Normality

We tested two main assumptions before proceeding to select the best SPSS procedure for the available data. The first assumption is that the dependent variables are approximately normally distributed. The second assumption is the data series does not contain significant outliers. First, we verified the normality assumption and then checked for any significant outliers. We used the numerical method of the Shapiro-Wilk normality test to test normality rather than the graphical method because the numerical method is more reliable.

The Shapiro-Wilk test null hypothesis is: the dependent variables are normally distributed (H0). The alternative hypothesis of the Shapiro-Wilk test is: the dependent variables are not normally distributed (H1). We accept the null hypothesis if we have asymptotic significance > 0.05. The results indicate that all the 20 dependent variables are not normally distributed with a confidence level of 99%. Therefore, we reject the null hypothesis. Alternatively, we accept the alternative hypothesis:

the dependent variables are not normally distributed (H0). The data series is negatively skewed with most of the respondent ratings between 3 and 5 and very few ratings in the 1 and 2 range.

The overall results indicate that all the 20 dependent variables are not normally distributed with a confidence level of 99%. Therefore, we reject the null hypothesis. Alternatively, we accept the alternative hypothesis: the dependent variables are not normally distributed (H0). The data series is negatively skewed with most of the respondent ratings between 3 and 5 and very few ratings in the 1 and 2 range.

Appendix B.3. Test for Significant Outliers

The outliers or extreme values may represent a danger for our proposed analysis because they affect the mean and standard deviation. These outliers are not due to data entry errors nor data, measurement of data collection errors. Data entry was electronic through the use of an electronic program known as SurveyMonkey (https://www.surveymonkey.com). Therefore the chances of data entry errors due to lack of attention, negligence, and tiredness were eliminated. Data collection was also done electronically through the participants filing the online questionnaires thus eliminating errors due to human mistakes or equipment malfunction. Also, data transfer from SurveyMonkey to SPSS was not manual. Thus, eliminating any form of data transfer errors. Therefore, the outliers we were dealing with are genuine non-typical and unusual values in the population.

We had two basic solutions to deal with genuine outliers. The first option was to remove the outlier from the data series. The second option was to keep the outliers in the data series. We realized that removing the outliers would remove all the assessment results of respondent ID number 1. It would also mainly impact on the results of some respondents, for instance, respondent ID number 3 and 11. To remove these extreme values would mean removing half of the dataset. Therefore, we decided to keep the outliers.

Appendix B.4. Chi-Square Test for the Goodness-of-Fit

The Chi-Square test for the goodness-of-fit test for cooperation was conducted on ten variables related to cooperation. The expected values for the theoretical distribution selected were "all categories equal." Though the number of respondents was 35, there were some missing values which led to different theoretical distributions. The minimum expected frequency of 7 was not met because there was no variable whose results had a response from all the five Likert scale options. The expected cell frequency is 8.8 per category, where responses fell into four categories and 11.7 per category, where responses fell into three categories.

The Chi-Square test for the goodness-of-fit test for team interdependence was conducted on ten variables related to team interdependence. The expected values for the theoretical distribution selected were "all categories equal." Though the number of respondents was 35, there were some missing values which led to different theoretical distributions. The expected cell frequency is 7.0 per category, where responses fell into five categories; 8.8 per category, where responses fell into four categories and 11.7 per category, where responses fell into three categories.

References

1. Government of Kenya (GoK). *National Water Service Strategy (NWSS)*; Irrigation, M.W., Ed.; GoK: Nairobi, Kenya, 2007. Available online: http://urbanlex.unhabitat.org/sites/default/files/urbanlex/nwss_0.pdf (accessed on 26 December 2018).
2. Water Services Regulatory Board (WASREB). A Performance Report of Kenya's Water Services Sector 2015/16 and 2016/17. In *Impact, Issue No. 10*; Government of Kenya: Nairobi, Kenya, 2018; p. 92. Available online: https://wasreb.go.ke/downloads/WASREB_IMPACT_Issue10_FINAL.pdf (accessed on 25 December 2018).
3. Water Services Regulatory Board (WASREB). *Water Services Regulatory Board (WASREB) Strategic Plan 2018–2022*; Government of Kenya: Nairobi, Kenya, 2018. Available online: https://wasreb.go.ke/strategic-plan/ (accessed on 25 December 2018).

4.　Pahl-Wostl, C. A conceptual framework for analysing adaptive capacity and multi-level learning processes in resource governance regimes. *Glob. Environ. Chang.* **2009**, *19*, 354–365. [CrossRef]

5.　Biswas, A.K.; Tortajada, C. Future water governance: Problems and perspectives. *Int. J. Water Resour. Dev.* **2010**, *26*, 129–139. [CrossRef]

6.　Rogers, P.; Hall, A.W. *Effective Water Governance*; Global Water Partnership Stockholm: Stockholm, Sweden, 2003; Volume 7. Available online: https://www.gwp.org/globalassets/global/gwp-cee_files/regional/governance-2003.pdf (accessed on 21 January 2019).

7.　Government of Kenya. *2017/2018 National Water Resources Situation Report*; Authority, W.R., Ed.; Government of Kenya: Nairobi, Kenya, 2018. Available online: https://www.wra.go.ke/annual-reports/ (accessed on 28 December 2018).

8.　Worldometers.info. Kenya Population. 2018. Available online: http://www.worldometers.info/world-population/kenya-population/ (accessed on 28 December 2018).

9.　UNDP. *Sustainable Development Goals*; UNDP: New York, NY, USA, 2015; Available online: http://www.undp.org/content/dam/undp/library/corporate/brochure/SDGs_Booklet_Web_En.pdf (accessed on 28 December 2018).

10.　Government of Kenya. *The Water Act 2016*; Government of Kenya: Nairobi, Kenya, 2016. Available online: http://kenyalaw.org/kl/fileadmin/pdfdownloads/Acts/WaterAct_No43of2016.pdf (accessed on 18 January 2019).

11.　Business Dictionary. Cooperation. 2017. Available online: http://www.businessdictionary.com/definition/cooperation.html (accessed on 4 December 2017).

12.　Kenya Commission on Revenue Allocation. *Kenya: County Factsheets*; Government of Kenya: Nairobi, Kenya, 2013. Available online: http://www.crakenya.org/wp-content/uploads/2013/07/CRA-County-Fact-sheets-2nd-Edition-June-2013.pdf (accessed on 22 October 2017).

13.　Government of Kenya. *Sessional Paper No. 1 of 1999 on National Policy on Water Resources Management and Development*; Government of Kenya: Nairobi, Kenya, 1999. Available online: https://www.ircwash.org/sites/default/files/824-KE99-18207-1.pdf (accessed on 18 January 2019).

14.　Government of Kenya. WATER SECTOR STRATREGIC PLAN (WSSP) 2009–2014. Available online: http://www.hakijamii.com/wp-content/uploads/2016/05/Water-Sector-Strategy-Plan.pdf (accessed on 9 September 2018).

15.　Government of Kenya. *The constitution of Kenya 2010*; Government of Kenya: Nairobi, Kenya, 2010. Available online: http://www.kenyalaw.org:8181/exist/kenyalex/actview.xql?actid=Const2010 (accessed on 4 September 2018).

16.　Onencan, A.M.; Van de Walle, B.; Enserink, B.; Chelang'a, J.; Kulei, F. WeShareIt Game: Strategic foresight for climate-change induced disaster risk reduction. *Procedia Eng.* **2016**, *159*, 307–315. [CrossRef]

17.　Onencan, A.M.; Van de Walle, B. Designing Disaster Diplomacy in the Context of a Climate Change Water Game. In Proceedings of the 48th International Simulation and Gaming Association Conference (ISAGA 2017) on Simulation Games for Sustainable Cities and Smart Infrastructures, Delft, The Netherlands, 10–14 July 2017; pp. 43–57.

18.　Onencan, A.M.; Van de Walle, B. Equitable and Reasonable Utilization: Reconstructing the Nile Basin Water Allocation Dialogue. *Water* **2018**, *10*, 707. [CrossRef]

19.　Onencan, A.M.; Van de Walle, B. From Paris Agreement to Action: Enhancing Climate Change Familiarity and Situation Awareness. *Sustainability* **2018**, *10*, 1929. [CrossRef]

20.　Onencan, A.M.; Kortmann, R.; Kulei, F.; Enserin, B. MAFURIKO: Design of Nzoia basin location based flood game. *Procedia Eng.* **2016**, *159*, 133–140. [CrossRef]

21.　Onencan, A.M.; 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]

22.　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.

23.　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; pp. 161–172.

24.　Onencan, A.; Meesters, K.; Van de Walle, B. Methodology for participatory gis risk mapping and citizen science for solotvyno salt mines. *Remote Sens.* **2018**, *10*, 1828. [CrossRef]

25. Onencan, A.; Enserink, B.; Van de Walle, B. A Study of Trust and Cooperation in the Nzoia River Basin Using a Water Policy Game. *Sustainability* **2018**, *10*, 4678. [CrossRef]

26. Enserink, B.; Onencan, A. Nile Basin Scenario Construction. In *IAIA17: Impact Assessment's Contribution to the Global Efforts in Addressing Climate Change: 37th Annual Conference of the International Association for Impact Assessment*; IAIA: Fargo, ND, USA, 2017; pp. 1–6.

27. Otieno, E. Kenya: Water Crisis in Migori Town After Taps Run Dry for Five Months. *Daily Nation* **2018**. Available online: https://allafrica.com/stories/201804080037.html (accessed on 8 May 2018).

28. UNECE, W. *Convention on the Protection and Use of Transboundary Watercourses and International Lakes*; United Nations Economic Commission for Europe Geneva: Genève, Switzerland, 1992.

29. UN-Water. *Integrated Monitoring Guide for SDG 6: Targets and Global Indicators*; United Nations: Geneva, Switzerland, 2016.

30. Hall, D.G. The Mouse Game and its Effects on Team Interdependence. *Dev. Bus. Simul. Exp. Learn.* **2014**, *38*, 3.

31. Depping, A.E.; Mandryk, R.L.; Johanson, C.; Bowey, J.T.; Thomson, S.C. Trust Me: Social Games are Better than Social Icebreakers at Building Trust. In Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play, Melbourne, Australia, 28–31 October 2018; pp. 116–129.

32. Wohn, D.Y.; Lampe, C.; Wash, R.; Ellison, N.; Vitak, J. The S in social network games: Initiating, maintaining, and enhancing relationships. In Proceedings of the 2011 44th Hawaii International Conference on System Sciences (HICSS), Kauai, HI, USA, 4–7 January 2011; pp. 1–10.

33. Depping, A.E.; Mandryk, R.L. Cooperation and Interdependence: How Multiplayer Games Increase Social Closeness. In Proceedings of the Annual Symposium on Computer-Human Interaction in Play, Amsterdam, The Netherlands, 15–18 October 2017.

34. Onencan, A.M. TU Delft Serious Game Elevates Nzoia. Delft, The Netherlands. 2017, p. 12. Available online: https://repository.tudelft.nl/islandora/object/uuid%3A91914ae3-453b-458a-ab58-555789a4533c (accessed on 5 June 2018).

35. UN-Water. *Step-by-Step Monitoring Methodology for Indicator 6.5.2*; UN-Water: Geneva, Switzerland, 2017.

36. UN-Water. *Draft Step-by-Step Monitoring Methodology for Indicator 6.5.2*; UN-Water: Geneva, Switzerland, 2016.

37. Hussein, H.; Menga, F.; Greco, F. Monitoring Transboundary Water Cooperation in SDG 6.5. 2: How a Critical Hydropolitics Approach Can Spot Inequitable Outcomes. *Sustainability* **2018**, *10*, 3640. [CrossRef]

38. Wolf, A.T. Shared waters: Conflict and cooperation. *Annu. Rev. Environ. Resour.* **2007**, *32*, 241–269. [CrossRef]

39. Government of Kenya. *Water Resources Management Authority (WRMA) National Master Plan 2030*; Government of Kenya: Nairobi, Kenya, 2013; p. 219.

40. McCracken, M.; Meyer, C. Monitoring of transboundary water cooperation: Review of Sustainable Development Goal Indicator 6.5. 2 methodology. *J. Hydrol.* **2018**, *563*, 1–12. [CrossRef]

41. Grey, D.; Sadoff, C.; Connors, G. *Effective Cooperation on Transboundary Waters*; Claudia: Connors, Genevieve, 2016.

42. Cascão, A.E. Changing power relations in the Nile river basin: Unilateralism vs. cooperation? *Water Altern.* **2009**, *2*, 245–268.

43. Hussein, H.; Grandi, M. Dynamic political contexts and power asymmetries: The cases of the Blue Nile and the Yarmouk Rivers. *Int. Environ. Agreem. Politics, Law Econ.* **2017**, *17*, 795–814. [CrossRef]

44. Hussein, H. The Guarani Aquifer System, highly present but not high profile: A hydropolitical analysis of transboundary groundwater governance. *Environ. Sci. Policy* **2018**, *83*, 54–62. [CrossRef]

45. Hussein, H. Whose 'reality'? Discourses and hydropolitics along the Yarmouk River. *Contemp. Levant* **2017**, *2*, 103–115. [CrossRef]

46. Hussein, H. *An Analysis of the Discourse of Water Scarcity and Hydropolitical Dynamics in the Case of Jordan*; University of East Anglia: Norwich, UK, 2016.

47. Selby, J. Cooperation, domination and colonisation: The Israeli-Palestinian joint water committee. *Water Altern.* **2013**, *6*, 1.

48. Selby, J. Dressing up domination as 'cooperation': The case of Israeli-Palestinian water relations. *Rev. Int. Stud.* **2003**, *29*, 121–138. [CrossRef]

49. Conker, A. *An Enhanced Notion of Power for Inter-State and Transnational Hydropolitics: An Analysis of Turkish-Syrian Water Relations and the Ilısu Dam Conflict between the Opponents and Proponents of the Dam*; University of East Anglia: Norwich, UK, 2014.

50. Warner, J.; Mirumachi, N.; Farnum, R.L.; Grandi, M.; Menga, F.; Zeitoun, M. Transboundary 'hydro-hegemony': 10 years later. *Wiley Interdiscip. Rev. Water* **2017**, *4*, e1242. [CrossRef]

51. Cox, R.W. Gramsci, hegemony and international relations: An essay in method. *Millennium* **1983**, *12*, 162–175. [CrossRef]

52. Julien, F. Hydropolitics is what societies make of it (or why we need a constructivist approach to the geopolitics of water). *Int. J. Sustain. Soc.* **2012**, *4*, 45–71. [CrossRef]

53. Menga, F. Hydropolis: Reinterpreting the polis in water politics. *Political Geogr.* **2017**, *60*, 100–109. [CrossRef]

54. Jägerskog, A. *Why States Cooperate over Shared Water: The Water Negotiations in the Jordan River Basin*; Linköping University Electronic Press: Linköping, Sweden, 2003.

55. Zeitoun, M.; Mirumachi, N. Transboundary water interaction I: Reconsidering conflict and cooperation. *Int. Environ. Agreem. Politics Law Econ.* **2008**, *8*, 297. [CrossRef]

56. Morgan, B.B., Jr.; Glickman, A.S.; Woodard, E.A.; Blaiwes, A.S.; Salas, E. *Measurement of Team Behaviors in a Navy Environment*; Battelle Columbus Labs Research Triangle Park NC: Orlando, FL, USA, 1986.

57. Onencan, A.M.; Enserink, B.; van de Walle, B. *Game Design Concept Report: Application of the We Share It Game Elements in Nzoia River Basin*; Delft University of Technology: Delft, The Netherlands, 2018; p. 53.

58. Barchiesi, S.; Sanchez, J.C.; Cross, K.; Pérez de Madrid, M.; Onencan, A.M. Adaptation planning: Views towards resilience and up-scaling success to enhance transboundary water governance. In *IUCN Environmental Policy and Law Paper No. 75*; IUCN Gland: Gland, Switzerland, 2014.

59. Wu, X.; Jeuland, M.; Sadoff, C.; Whittington, D. Interdependence in water resource development in the Ganges: An economic analysis. *Water Policy* **2013**, *15*, 89–108. [CrossRef]

60. De Chaisemartin, M. Measuring Transboundary Water Cooperation within the Framework of Agenda 2030: A Proposal for a Revision of SDG Indicator 6.5.2. *Soc. Sci. Res. Netw.* **2018**. [CrossRef]

61. UNECE Water Convention Secretariat. Reporting Under the Water Convention and Monitoring of Global SDG Indicator 6.5.2. In *Working Group on Integrated Water Resources Management*; Meeting, E., Ed.; UNECE: Brussels, Belgium, 2016.

62. Swain, A. The Nile River Basin Initiative: Too many cooks, too little broth. *SAIS Rev.* **2002**, *22*, 293–308. [CrossRef]

63. Zeitoun, M. *Power and Water in the Middle East: The Hidden Politics of the Palestinian-Israeli Water Conflict*; IB Tauris: London, UK, 2008.

64. Nile Basin Initiative (NBI). *State of the River Nile Basin Report*; Nile Basin Initiative: Entebbe, Uganda, 2012.

65. Garris, R.; Ahlers, R.; Driskell, J.E. Games, motivation, and learning: A research and practice model. *Simul. Gaming* **2002**, *33*, 441–467. [CrossRef]

66. Klein, G.A. *A Recognition-Primed Decision (RPD) Model of Rapid Decision Making*; Ablex Publishing Corporation New York: New York, NY, USA, 1993.

67. Sultana, F. An (Other) geographical critique of development and SDGs. *Dial. Hum. Geogr.* **2018**, *8*, 186–190. [CrossRef]

68. Lorenz, E.H. Neither Friends nor Strangers: Informal Networks of Subcontracting in French Industry. In *Trust: Making and Breaking Cooperative Relations*; Gambetta, D., Ed.; Basil Blackwell Ltd.: Cambridge, MA, USA, 1988; pp. 194–210.

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

70. Ortigara, A.; Kay, M.; Uhlenbrook, S. A review of the SDG 6 synthesis report 2018 from an education, training, and research perspective. *Water* **2018**, *10*, 1353. [CrossRef]

71. McCracken, M. *Measuring Transboundary Water Cooperation: Options for Sustainable Development Goal Target 6.5*; Global Water Partnership (GWP): Stockholm, Sweden, 2017.

72. Tarlock, D. Promoting effective water management cooperation among riparian nations. *Tech. Comm. Backgr. Paper* **2015**, 64. Available online: https://www.gwp.org/globalassets/global/toolbox/publications/background-papers/gwp_tec21_web.pdf (accessed on 21 January 2019).

73. Strategic Foresight Group. *Water Cooperation Quotient*; Strategic Foresight Group: Mumbai, India, 2017. Available online: https://www.strategicforesight.com/publication_pdf/Water%20Cooperation%20Quotient%202017.pdf (accessed on 30 December 2018).

74. Oregon State University, C.O.E. Ocean and Atmospheric Sciences. Transboundary Freshwater Dispute Database (TFDD) Basin at Risk (BAR)/International Water Intensity Scale Oregon State University, Program in Water Conflict Management and Transfomation: Nd. Available online: https://transboundarywaters.science.oregonstate.edu/content/basins-risk (accessed on 30 December 2018).

75. Sadoff, C.W.; Grey, D. Cooperation on international rivers: A continuum for securing and sharing benefits. *Water Int.* **2005**, *30*, 420–427. [CrossRef]

76. Sadoff, C.W.; Grey, D. Beyond the river: The benefits of cooperation on international rivers. *Water Policy* **2002**, *4*, 389–403. [CrossRef]