Beyond Educational Policy Making

With Agent-Based Simulation

Atsushi Yoshikawa[1(✉)] and Satoshi Takahashi[2]

[1] Tokyo Institute of Technology, Yokohama-shi, Kanagawa 226-8503, Japan
at_sushi_bar@dis.titech.ac.jp
[2] Tokyo University of Science, Chiyoda-ku, Tokyo 102-0071, Japan

Abstract. In recent years, formulation of educational policy has come to be based on data. That data, however, can turn out to be difficult to access, or mixed with so much noise interfering with education policy formulation, that it cannot be used directly for policy making. To address this issue, an increasing number of attempts to contribute to policy formulation have been made using agent-based simulation (ABS). In the majority of research, ABS is used in the ex post facto analysis of why educational policy has not been effective. In this paper, case studies show that by incorporating ABS into the policy formulation process, the risk of failure can be reduced. By illustrating the relationships between model level, stage of educational policy formulation and the output scenarios of ABS, it is possible to determine which types of risks can be reduced. This paper presents ABS description levels, and discusses risks that both can and cannot be expressed using ABS. We show two ways to use ABS for educational policy making by identifying risks that can be reduced and risks that cannot be dealt with by ABS.

Keywords: Educational policy making · Agent-based simulation · Real-world data · Modeling of educational policy making

1 Introduction

Education policy not only affects the lives of individuals in the future. It also has a major effect on society, where the human resources produced by that education system are employed. The importance of education policy is widely acknowledged, and there are numerous reports describing outcomes of education policy in various countries, both successful and less effective. However, there are few studies on educational policy that explain how the policy has been derived and why it was considered valid at the time it was conceived. An educational policy that was a Japanese version of the No Child Left Behind (NCLB) Act was implemented in Japan about 20 years ago. Called ‘yutori’ education, or ‘education with breathing space’, the results of this policy have been assessed as contributing to the lowering of academic ability in Japan [1]. Some predicted the failure of this policy from the outset [2]. Even though there has been some self-reflection regarding the failure of the policy, to date there has been no research into the reasons why such a policy was adopted. Educational policies are determined after discussion in meetings and conferences, in the course of which data from different

F. Koch et al. (Eds.): GEAR 2018, CCIS 999, pp. 112–130, 2019.
https://doi.org/10.1007/978-981-13-6936-0_12

studies are presented. If a policy leads to unanticipated results, an explanation should be sought. Educational policies that are less successful than envisaged are not limited to Japan. Educational policies related to the NCLB Act and the Perry Preschool Study were based on data and underwent examination but did not bring complete success [3, 4].

For this paper, we categorize precedents in agent-based simulation (ABS) related to educational policy as a method for formulating policy that offers a greater reduction in the risks of failure [5]. From that, we identify distinctive characteristics of ABS that can be linked with educational policy, and propose that introducing ABS into the process of formulating educational policy is likely to be effective.

2 Categorization of Educational Policy

2.1 Levels of Educational Policy

UNESCO states that the determination of educational policy (in the broad sense) involves three stages of (1) policy (referred to here in the narrow sense; hereafter, policy in this narrow sense will be referred to as 'policy measures' and in the broad sense it will be referred to as educational policy); (2) strategy, and (3) plans [6]. According to UNESCO's definition (here slightly revised), these are:

- *A <u>policy</u> establishes the main goals and priorities pursued by the government in matters of education – at the sector and sub-sector levels – with regard to specific aspects such as access, quality and teachers, or to a given issue or need.*
- *A <u>strategy</u> specifies how the policy goals are to be achieved.*
- *A <u>plan</u> defines the targets, activities to be implemented and the timeline, responsibilities and resources needed to realize the policy and strategy.*

Policy determines goals and priorities, typically determined by the external environment. In line with that determination, strategy devises methods for proceeding toward realization. Plans for concrete realization of those methods are, in short, the instantiation of policy in the form of targets and activities intended to achieve the strategy. UNESCO recommends that these three stages be carefully separated according to their function. Naturally, good policy measures do not appear full-blown from the very start, so it is envisaged that the four-part policy cycle of analysis–planning–implementation–evaluation will undergo numerous iterations. A look at this cycle will show, however, that though it is effective for improving an educational policy that has been decided on, it offers no guidance for determining new educational policy from the ground up. UNESCO itself further finds that effective educational policy has to be [6]:

- *Built on evidence*
- *Politically feasible*
- *Financially realistic*
- *Agreed to by the government and relevant stakeholders.*

While imposing these strict constraints, they state that the policy cycle is preceded by this vision:

Step 0 (Vision): Before the start of a policy cycle, a strategic intent, often called a "vision", is formed. For instance, once a political party wins a majority of seats in parliament and forms a government, they denote their strategic intent for education, which, for instance, may be: "Increase participation of youth from lower socio-economic backgrounds in tertiary education."

This is a weak point in the process of policy formulation.

2.2 Areas Addressed by Educational Policy

UNESCO delineates the following three levels of macro, meso, and micro as the areas addressed by educational policy [7].

(a) Macro: region/state/national/international
(b) Meso: institution-wide
(c) Micro: individual user actions.

The evidence required to formulate policy differs depending on the area addressed, but this does not mean that the necessary evidence for the various areas exists. For instance, Japan's 'education with breathing space' system was formed by policy measures of the national government, and thus consists of macro policy measures. The system is set up so that matters are then decided in various derivative areas of policy with measures branching out from each level. These measures range from national curriculum standards to textbooks. Teacher training is the responsibility of the boards of education at the local government level [8]. A specific educational method was even devised for the classroom context, known as the integrated study period. This means that the process has been determined by policy and strategy. Following Japanese tradition, the plans are left up to the schools [9]. The shift toward 'education with breathing space' was probably carefully examined because there were data from subjective assessments conducted in the course of students' studies for entrance exam competitions, and qualitative data in the form of observations voiced by people involved in classroom education. However, no data exist on the time taken in teacher training. If data were available, it would be possible that some teachers experience unanticipated difficulties in adequately applying the approach to education [10]. These data would be part of planning, and may be difficult to obtain or does not exist. The result is that the information needed to enact educational policy is incomplete, and this is conceivably the reason that conditions emerged that were unanticipated by experts. This illustrates that when data supporting the level of the educational policy do not exist, unanticipated conditions tend to occur.

2.3 Data for Making Decisions About Educational Policy

Unless data for each of the policy–strategy–plan levels exist, the formulation of educational policy will be insufficiently grounded. However, it may also be unrealistic to obtain the full range of these data.

To formulate educational policy, real-world data from factual investigations of actual circumstances and other such sources are used. However, these real-world data may not be as useful as expected in informing educational policy, since such data may contain distortions due to various circumstances. For instance, there may be data showing that 100% of schools are equipped with personal computers. However, even if this is the case, there may be schools where computers are locked away in storage so that students cannot damage the expensive equipment, or where the computers are so old that they do not function properly and are of little use. In these cases, the computer adoption rate will be high, and if policy is made in the expectation that courses will involve the use of computers, then that policy will be out of touch with reality.

The real-world data used in formulating educational policy have the following limitations.

(a) Ethical considerations prevent control experiments from being performed, so the data do not contribute directly to policy aims.
(b) Real-world data are difficult to interpret because phenomena are overly complex.
(c) Long-term influences exist in data.
(d) Data cannot be acquired in the first place.

These are matters to be aware of at the stage of creating an educational policy. The weakness, however, is that measures to resolve the problems are difficult to come by. Item (d) is a problem that may not be possible to address, but the first three items are worth considering further.

Ethical Considerations Prevent Control Experiments from Being Performed, so the Data Do Not Contribute Directly to Policy Aims. When implementing policy measures, the scientific approach is to conduct experiments with the participants divided into control and experimental groups. Comparison of results from the two groups confirms the validity of the proposed method. These are known as control experiments. The field of education, however, faces the issue of equality of educational opportunity. This means that using control groups is not possible. At the same time, implementing an educational policy that cannot be expected to be effective is not ethically acceptable. Consequently, policy measures cannot be confirmed experimentally, even in part.

Real-World Data Are Difficult to Interpret Because Phenomena Are Overly Complex. The effectiveness of education is influenced not only by teachers and instructional materials, but also by the characteristics of the group of learners who receive the same education and by the environment. These factors exist across a range from the individual learner level to the macro level. At the individual learner level, factors may include the age and gender of the learner, the family structure, and so on. At the middle level, they may include the make-up of classes at the school or other educational institution, and the curriculum design. At the macro level, they may include the configuration of the school district, the placement of the educational institutions, and the system for advancing students to the next level of education. As this illustrates, the factors are interrelated at multiple levels in such complex ways that it is difficult to interpret acquired data.

Long-Term Influences Exist in Data. Developing human resources takes a long time. For that reason, there will not be just one educational policy that exerts an influence over that development period, and it is difficult to assess a policy and associated policy measures as a stand-alone unit. For instance, grades in lower secondary school mathematics are sometimes affected by insufficient acquisition of Japanese language competence in primary school [11]. Furthermore, a single educational policy can affect not only short-term changes in academic ability, but also long-term outcomes. There is the extremely rare case of the study of long-term influence in the Perry Preschool Study. This study of an early childhood education program in the United States, conducted from 1962 to 1967, carried out a continuous analysis of the influence of the educational program. The long-term influence continues to be studied even today [4]. Cases like this, however, are extremely unusual.

There is progress in addressing the deficiencies of data used for educational policy. With regard to the issue of ethical considerations, measures have been taken, for example, to use alternative experimental or quasi-experimental designs (e.g. crossover trials) and establish special districts related to education, e.g. charter schools in the United States [12]. An example of a quasi-experiment includes the assessment of educational subjects related to Title I schools, which makes use of differences between schools in their circumstances [13]. These are measures devised in the actual context of education in practice, and in that sense can be considered measures devised out of awareness of the limits involved.

2.4 Weaknesses in Educational Policy

A summary of results indicates that the course of formulating educational policy is subject to the following three weaknesses:

(a) When formulating new educational policy, we cannot expect to find real-world data that would serve as a reference.
(b) When policy-making lacks data that support the level of educational policy concerned, unanticipated conditions are likely to occur.
(c) When formulating educational policy, the real-world data that are used will have certain limitations.

3 Application of ABS to Educational Policy

3.1 Potential of ABS as Data

Gilbert states that the three weaknesses of data presented for the purpose of framing educational policy (referred to in Sect. 2.3), could be resolved by the use of ABS. This can be inferred from the distinctive characteristics of ABS listed below [5].

(a) As a simulation, it can be used to try out any configuration.
(b) It is specialized in the handling of complex systems.
(c) As a simulation, it allows the time interval to be freely selected.

As a Simulation, ABS Can Be Used to Try Out Any Configuration. In ABS, a virtual world is built and the agent is made to act within that world to carry out the simulation. The designers can try out any configuration they like in the virtual world. This distinctive characteristic makes it possible to conduct control experiments that would be difficult in the real world.

ABS is Specialized in the Handling of Complex Systems. With this technique, people are modeled as individuals, called agents, who act autonomously, and the real world in which people act is modeled as a virtual world. The simulation is then carried out by having the agents interact with each other and with the virtual world. This approach is therefore suited to dealing with complex phenomena caused by the interaction of elements at education-related levels ranging from the micro to the macro.

As a Simulation, ABS Allows the Time Interval to Be Freely Selected. In other words, the length of time taken to carry out the simulation can be configured freely as the designers wish. For that reason, this approach allows examination of the long-term influence of educational policy by means of simulation.

Due to these distinctive characteristics of ABS, data can be provided for the purpose of formulating educational policy that could not be achieved with real-world data alone. ABS is model-based, so it is created on the basis of abstract models in education. In other words, it is created with a variety of educational theories as its basis. Its parameters are derived from real-world data. The ABS results themselves are not used extensively. Rather, the various scenarios generated by ABS are interpreted to allow examination of: (a) the process by which the results of educational policy come about, (b) impact from the educational policy that is formulated, (c) divergence from expected values, and (d) how policy created to improve the situation might actually end up making it worse instead.

3.2 Model Levels

Gilbert explains that, like educational policy, ABS has three levels [5]. These are the abstract, middle range, and facsimile levels.

An abstract model has the lowest level of resolution. The model is constructed out of a small number of parameters and interactions, with the aim of gaining a fundamental understanding of a social phenomenon. A classic example is Schelling's segregation model [14].

The middle range model has a slightly higher resolution than the abstract model. The component elements and parameters are markedly more numerous than in the abstract model. One example that can be cited is the research of Yano et al. [15].

The facsimile model is the model with the highest resolution. It aims to recreate a social system under specific circumstances. One example is the model of a specific classroom.

According to the conceptual approach of this categorization, real-world educational policy consists of policy, strategy, and plan stages. They correspond to the facsimile model since they are created with awareness of specific circumstances. The ABS, by contrast, contributes to the formulation of educational policy by the interpretation of scenarios. Therefore it is a model of at least the middle range or higher (Fig. 1).

Figure 1 shows a schematic representation of the relationship between educational policy and ABS. It shows that ABS, through interpretation of scenarios, can yield data at the plan level of educational policy.

4 Cases of ABS Application to Educational Policy

4.1 Introduction of the Case Examples

Educational policy is formulated in a variety of fields. Here, therefore, it will be placed under the three classifications of macro, meso, and micro named in Sect. 2.2, and real-world educational policy and ABS will be viewed on that basis.

Model at Macro Level. Macro examples that could be cited include the NCLB Act, school system model, school district scope, and guidelines for establishing schools. The case of the NCLB Act is introduced here.

Case Example: About the No Child Left Behind Act. The NCLB Act was an educational policy in the United States that was passed into law under the administration of the country's 43rd president, George Walker Bush, in 2002 [16]. NCLB calls for student learning progress goals (adequate yearly progress: AYP) to be set state by state, and reading, math, and science are the areas for assessment. Progress toward goals is assessed by tests. Schools that do not achieve their goals are negatively viewed and required to take corrective action. The NCLB Act requires separate reports to be submitted for students of different income, race, disabilities, and English language learners. The purpose is to correct disparities in academic ability resulting from these factors, a distinctive characteristic of NCLB.

The US Department of Education has reported that various study results show links to improvements in students' academic ability [17]. For instance, the study results from the National Assessment of Educational Progress (NAEP) (https://nces.ed.gov/nationsreportcard/) are reported as follows:

> *For America's nine-year-olds in reading, more progress was made in five years than in the previous 28 combined.*
> *America's nine-year-olds posted the best scores in reading (since 1971) and math (since 1973) in the history of the report. America's 13-year-olds earned the highest math scores the test ever recorded.*
> *Reading and math scores for African American and Hispanic nine-year-olds reached an all-time high.*
> *Math scores for African American and Hispanic 13-year-olds reached an all-time high.*
> *Achievement gaps in reading and math between white and African American nine-year-olds and between white and Hispanic nine-year-olds are at an all-time low.*

However, various abuses have occurred, and in 2012, President Barack Obama, the 44th US president, decided to grant waivers [17]. 'Abuses' refer to the occurrence of conditions that had not been envisaged at the start of the plan. For instance, schools that were designated as 'failing' saw a decline in the morale of school staff and resignation

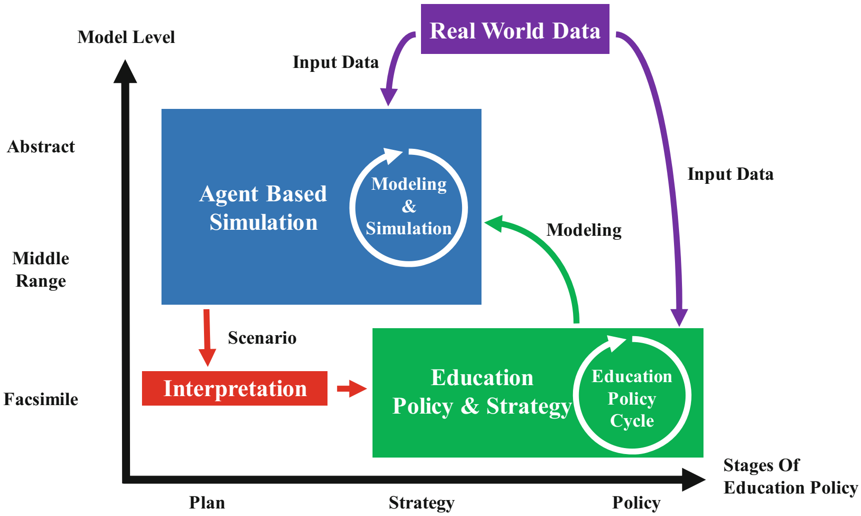

Fig. 1. Relationship between educational policy and agent-based simulation

rates increased [18]. In addition, since AYP assessment covers reading, math, and science, there was a tendency for other subjects not assessed to be neglected [19]. At the extreme, there were cases in which teachers of subjects other than those tested were fired and replaced by part-time instructors [20]. Schools that did not meet the prescribed levels were designated 'failing' schools. At schools that were designated as failing for two successive years, students were allowed to transfer to other schools – and conditions calling for transfer did actually occur.

Case Example of ABS Application: Causal Analysis of the NCLB Act by ABS. Sklar et al. suggest SimED as an example of ABS for use in examining educational policy [21]. In the SimED world, there are agents in the four roles of student, teacher, principal, and superintendent. These agents move around in a virtual world made up of the three levels of classroom, schoolhouse, and school district. The students have such variables as ability, motivation, emotion, belief set, satisfaction, family income, mobility, and performance. Using ABS with this kind of middle-range framework, Sklar et al. analyzed scenarios that could occur over a 20-year period following introduction of the NCLB. The introduction of a system of small class sizes (a 'STAR-like scenario' [22]) indicated the possibility that the percentage of students transferring to other schools could be lowered. This is primarily an ex post factor search for causes after a problem has occurred (in this case, scenario analysis). At the same time, however, Sklar et al. found that smaller class sizes would be a solution for the NCLB (Fig. 2).

Model at Meso Level. Meso level examples include curriculum design and class placement methods. The case of class size design will be introduced here.

Case Example: About Class Size Design. Class size design has been a matter of debate for many years, and large-scale experiments to verify the effectiveness of small class sizes have been carried out. For example, there is the Student Teacher Achievement Ratio (STAR) project in the state of Tennessee in the United States [19]. The STAR project tracked the changes in academic ability that were caused by class size in the four-year period from kindergarten through to third grade. The experiment was carried out with the following three class types: small classes (13–17 students per teacher), regular classes (22–25 students per teacher) and regular classes (22–25 students) with a full-time teacher's aide.

Results from the STAR project have been analyzed by different researchers, with contradictory findings. Konstantopoulos surveyed academic ability in mathematics and reading in relation to small classes at multiple schools, and found both cases of schools where academic ability rose and cases where it fell [23]. The situation with regular classes with a full-time teacher's aide was similar: the results showing instances where academic ability rose and other examples of where it fell.

There was also a five-year pilot project in the state of Wisconsin that started in the 1996/1997 academic year. Known as the Student Achievement Guarantee in Education (SAGE) project, this involved intervention experiments with measures such as lowering the number of students per teacher at participating schools to 15 [24]. The data collected were analyzed using regression analysis and hierarchical linear models (HLM), producing the conclusion that a small class size is effective in raising students' academic ability. The approach was particularly effective for students of African-American descent, indicating the possibility that having small class sizes could become an effective means of offsetting the difference in academic ability between African-American and white students.

Based on research to date, a conclusion on whether small class sizes are effective or not has yet to be reached.

Case Example of ABS Application: Analysis of Class Size Design by ABS. Yano et al. proposed an ABS model to use in analyzing class size [15]. The model contains agents in the two roles of student and teacher. These agents move and act in the same world, but the students and teachers belong to different schools, either primary, lower secondary, or upper secondary school. Multiple schools exist, and when students advance to the next level, they go to nearby schools. Students are characterized by variables such as academic ability, learning strategies, and academic motivation. Learning strategies follow three styles: teacher-driven, mutual teaching, and motivation-dependent. The students therefore proceed with their learning activities in accordance with their learning strategies, teaching each other, being educated by their teacher, and so on.

Yano et al. examined the scenarios that would be generated by where teacher assignments were concentrated in primary school, lower secondary school, or upper secondary school, at which small class sizes were realized. The results indicated that improvements in academic ability could be obtained among students at the lower levels of academic ability using the teacher-driven and mutual teaching styles of learning

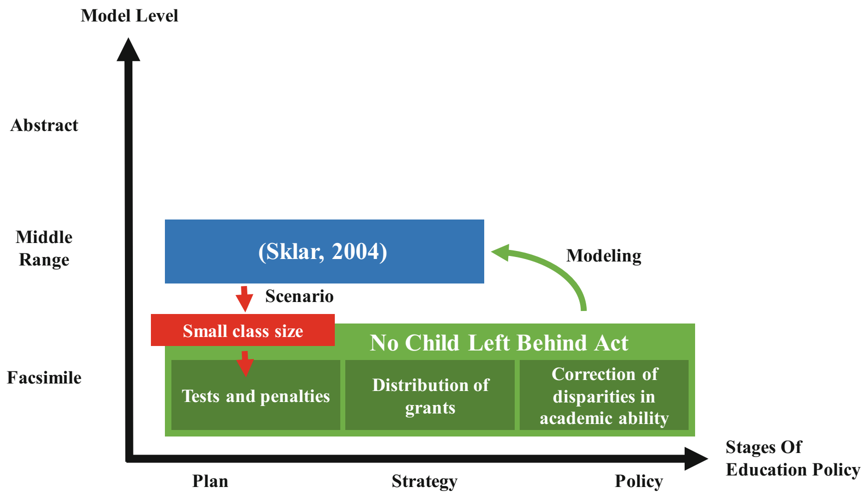

Fig. 2. Positioning the answers yielded by NCLB and SimED

when additional teacher agents were assigned to the primary schools. With the motivation-dependent style, however, these results were not observed. This indicates that results will differ according to student learning strategies. These results concur with the research of Konstantopoulos and Molnar [23, 24].

Interpreting these results, we find that having a system of small class sizes is given in educational policy, and that strategies and plans are positioned as practices at the educational site. In ABS, on the other hand, the model is positioned to place emphasis on the plan. This relationship is shown in Fig. 3.

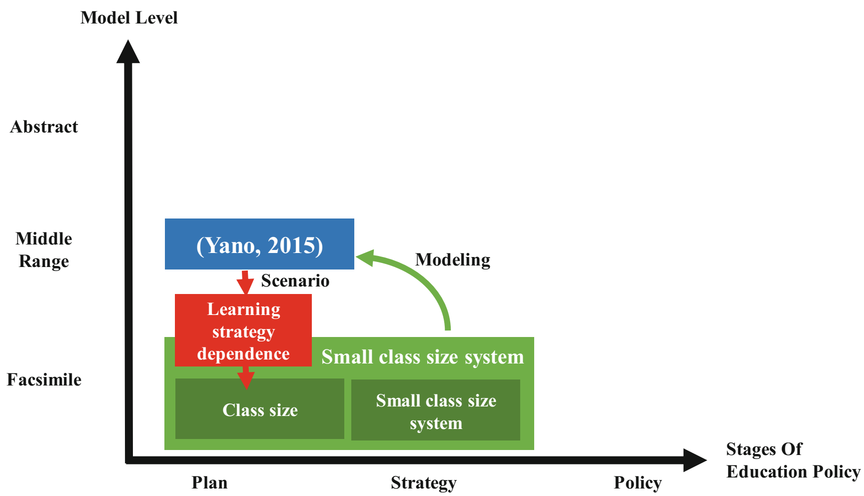

Fig. 3. Small class size in relation to solutions derived by the model of Yano et al.

Model at Micro Level. Micro level examples include design methods for collaborative and cooperative learning. Strict distinctions between collaborative and cooperative learning have been pointed out. Here, however, these are positioned as attempts to heighten learning effectiveness by having students influence each other through group learning [25].

Case Example: Collaborative Learning. At public schools in the United States during the 1970s, measures to eliminate racial segregation were implemented. The mistrust between whites and African-Americans was deeply rooted, however, and frequent clashes occurred. Aronson proposed the jigsaw method to build relationships of mutual trust during the primary education phase, when mistrust is relatively small [26]. In the jigsaw method, students are placed in small groups of five to six and members of each group are further placed in separate expert groups. The goal of the small groups is for each one to solve a problem together. The students are first separately given information in their respective expert groups. They are called on to assimilate that information and carry it back to their small groups. Next, the students are called on to integrate their pieces of information and solve their assigned problems. The problems are designed to yield solutions when the various pieces of information are well integrated. The aim is to have cooperative relationships form naturally among the students.

Aronson finds that use of the Jigsaw method leads to the formation of cooperative relationships as a result of learning: (1) that within their groups, nobody is able to learn without the help of others; and (2) that the members of the groups are able to make contributions that are uniquely their own and are essential.

Analysis of Collaborative and Cooperative Learning by ABS. Here we introduce the ABS approaches of Kuniyoshi et al. and Spoelstra et al. Kuniyoshi et al. examined the effectiveness of learning in terms of learners' academic ability, the structure of their instructional materials, and their collaborative relationships [27]. Spoelstra et al. examined the impact on group learning results from diversity of capabilities within groups, group sizes, and rewards to the groups [28].

Case example 1: Model of knowledge teaching on a complex doubly structural model taking academic ability, instructional material structure, and collaborative relationships into account

Kuniyoshi et al. did not use the jigsaw method in their research, but have modeled the inside of the Japanese classroom. Through this approach they have investigated the extent of influence that learners' individual academic ability, their instructional material structure, and their collaborative relationships have on learning effectiveness [27]. In the Kuniyoshi et al. model, agents use a 'complex doubly structural network' that models the structure of the learners' understanding of their instructional materials (knowledge network structure) and their collaborative relationships with each other (social network structure) [29]. Simulations are then carried out on the academic ability of the learners, the structure of knowledge about the instructional materials that they study, the teaching strategies of instructional staff, and the influence of seating arrangements on learning effectiveness. This yielded the following results:

(a) *When different teaching strategies, seating arrangements, and collaborative learning are used, learning effects vary.*

(b) *Group-style collaborative learning with a dispersed seating arrangement has the best effect on learning, and this would be followed by the method with the highest answer rate.*
(c) *An ability-based class has a negative effect on collaborative learning because diversity is reduced, and learner homogeneity decreases the collaborative effect.*
(d) *If teaching occurs only once for one knowledge item, learners may fall behind. Reviews should be conducted more frequently to facilitate the anchoring of the knowledge in a class.*

Case example 2: Diversity of group composition is problem-solving capability
Spoelstra et al. conducted analysis of learning effects in Student Teams Achievement Divisions (STAD) using ABS [28, 30]. The flow of learning in STAD is as follows [28]:

(a) *Teacher presentations: The initial phase of the learning process in which a teacher explains the concept to be acquired;*
(b) *Student teamwork or individual work: The phase in which activities designed to facilitate learning are undertaken by one or more students, working alone or in groups;*
(c) *Quizzes: The phase in which the teacher evaluates the progress made by each student;*
(d) *Individual improvement: The phase in which individuals receive recognition (from the teacher and/or their peers) for any progress they have made; and, optionally,*
(e) *Team recognition: The phase in which teams are ranked and prizes (or some other form of recognition) are bestowed upon team members. This phase is only relevant when the 'cooperative goal structure' is in place and students are working in teams.*

The models used by Spoelstra et al. have student agents. Student agents have state variables such as ability, emotion, motivation, and zone. They engage in learning in accordance with the STAD model and obtain rewards commensurate with their academic performance.

Spoelstra et al. used this ABS model to analyze diversity in the levels of capability in the group, group size, and rewards to the group, and how these influence learning effectiveness in STAD. This analysis indicated that diversity, group size, and rewards to the group exert influences on group learning effectiveness. In a particularly startling result, Spoelstra et al. emphasized the possibility that reward might be effective for a group made up entirely of people with low levels of ability and if the reward was of an appropriate size.

The relationship between educational policy and the Kuniyoshi ABS is shown in Fig. 4. There is less of a correspondence here than was shown in the macro and meso cases. However, what Kuniyoshi refers to as the comprehension of instructional materials is something the jigsaw expert group learned well, and if collaborative relationships are thought of as small jigsaw groups, then Kuniyoshi's ABS data could be interpreted as identifying further success factors contributing to the effectiveness of the jigsaw approach (Fig. 5).

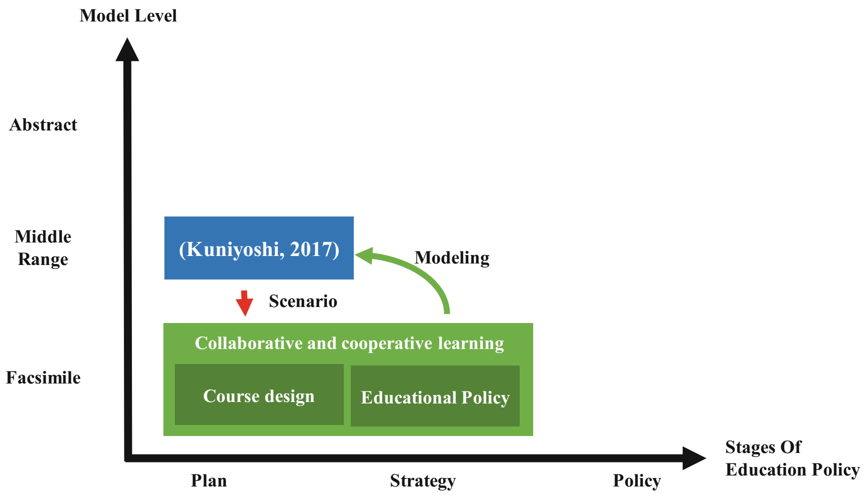

Fig. 4. Relationship between collaborative learning and the explanation of the Kuniyoshi et al. model.

4.2 Consistency of Educational Policy with ABS

The simulation results of ABS, representing the actuality of educational policy, can explain the results of educational policy implementation. However, the interpretation of ABS results may intrude because the model providing the basis for ABS is at a higher level of abstraction. This is because the data used for educational policy and educational structure provide the source for ABS. The data that are necessary in the formulation of educational policy can be configured as parameters in ABS, and the scenario will change depending on that configuration. Consequently, the particular scenario that is adopted will change the results that are obtained. Furthermore, even for just a single scenario, the simulation can be carried out for an enormous number of iterations so that it can find unexpected risks. This approach is able to perceive the pitfalls that lie in the path of educational policy. Arai et al. used ABS to derive the decline in academic ability that had not been envisaged when ‘education with breathing space’ was being designed [31].

Whether the target of the educational policy is at the macro, meso, or micro level, the positioning with regard to educational policy and ABS is maintained in the relationship shown in Fig. 1. When it comes to the interpretation of ABS results, however, depending on the ABS model, it may: (a) relate to specific policy measures at the plan level, or (b) turn out to be capable of explaining the reason why an educational policy brought about the results it did. The case examples show that ABS can be used in these two orientations.

5 Change in the Course of Educational Policy Formulation Due to ABS

In Sect. 2, the following three areas were identified as weaknesses occurring in the course of formulating educational policy:

(a) When formulating new educational policy, we cannot expect to find real-world data that would serve as a reference.
(b) When policy-making lacks data that support the level of educational policy concerned, unanticipated conditions are likely to occur.
(c) When formulating educational policy, the real-world data that are used will have certain limitations.

In Sects. 3 and 4, the potential of ABS to make up for these weaknesses was discussed. Based on the case examples, ABS was found to have two main capabilities:

(a) To explain the reason why an educational policy brought about the results it did;
(b) To examine specific policy measures at the plan level.

These findings will be placed into organized form as a method for using ABS in the course of formulating educational policy.

5.1 Application of ABS in the Course of Formulating Educational Policy

Examination of the process of educational policy formulation when that process is underway will produce something similar to Fig. 5. With regard to the first capability, the formulation of educational policy means that the ABS model can be structured on that basis. The results of that simulation can then be crosschecked with the data generated by implementation of the educational policy. Interpreting both sets of

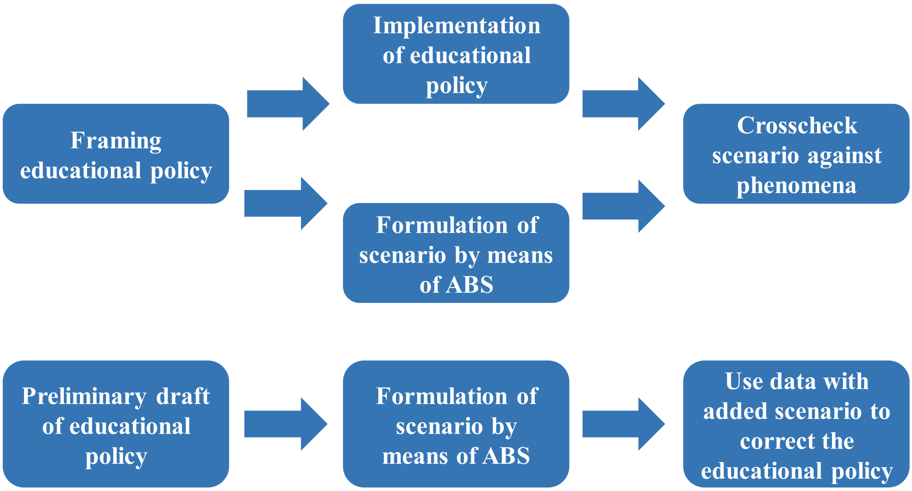

Fig. 5. Two pathways for introduction of ABS to educational policy

findings, makes it possible to explain the results. With regard to the second capability, the process of framing policy can be formulated as a scenario in ABS based on the preliminary draft of the policy. The scenarios generated by ABS are checked for unexpected risks. If necessary, the formulation of the scenario can also be repeated in ABS and corrections made.

In this way, methods for using ABS can be envisaged as Usage Method (1): Causal Analysis by ABS, and Usage Method (2): Identifying Scenarios by ABS. By employing ABS, causal analysis of educational policy can be carried out rapidly even after the policy has been formulated. In addition, formulating a scenario at the preliminary draft stage of a policy, can enable correction of the preliminary draft of a policy by incorporating matters that policy-makers could not have foreseen. Scenarios can be examined to determine whether they fall within what was envisaged, and are within the acceptable limits of risk.

Specifically, Usage Method (1) applies ABS to causal analysis of the reasons why education is in its current state, and of the mechanism by which an implemented educational policy led to its effect, or did not lead to an effect. For instance, verification of the effectiveness of a small class size also indicated the possibility that the variations in academic ability within a class are related to improvement in the academic ability of the class as a whole [32]. In this case, it was reasoned that exchange between students with higher academic ability and those with lesser abilities in a small class heightened academic motivation on both sides. By applying ABS in this situation, it is possible to analyze in greater detail situations such as how students with different levels of academic ability help each other.

In Usage Method (2), ABS is used for the examination of possible ways in which the preliminary draft of an educational policy can lead to results. By using ABS, scenarios that could be brought about by the educational policy can be identified, and on that basis, the educational policy improved and implemented. In connection with the NCLB Act, for instance, it is possible to analyze the types of scenarios that could occur in future as a result of state-by-state differences in fiscal resources, the economic status of residents, racial composition, and academic ability. Specifically, this examines the combination of conditions that lead to failure of a policy or a rise in student dropout rates. By identifying scenarios like these, scenarios of policy failure are discovered and measures to prevent that failure can be examined.

There are areas that ABS cannot handle such as rigorous forecasting and producing numerical values. ABS is applied to complex systems, and is sensitive to initial conditions. A small difference in conditions can yield great differences in results. In connection with the examination of small class sizes, for instance, this approach is not able to deciding exactly how many students should be in a class. It is also unable to forecast the specific quantitative effects of the NCLB Act.

5.2 Conditions for Application of ABS

ABS must be applied to phenomena that it is able to model. Considering the cases of model application in education analyzed by Yoshikawa et al., models can be divided into three categories: models based on mathematical, logical, and structural descriptions [33].

Models based on mathematical description are models that express the relationship between their component elements using numerical formulae. Models based on logical descriptions are models that express the logical relationship of their component elements. Logical relationships are expressed by figures and text that supplement these figures. Models based on structural descriptions refer to models that decompose their component elements and express the content of those elements as well as their structure and categorization. Structure and content are expressed by figures and text that supplement those figures (Fig. 6).

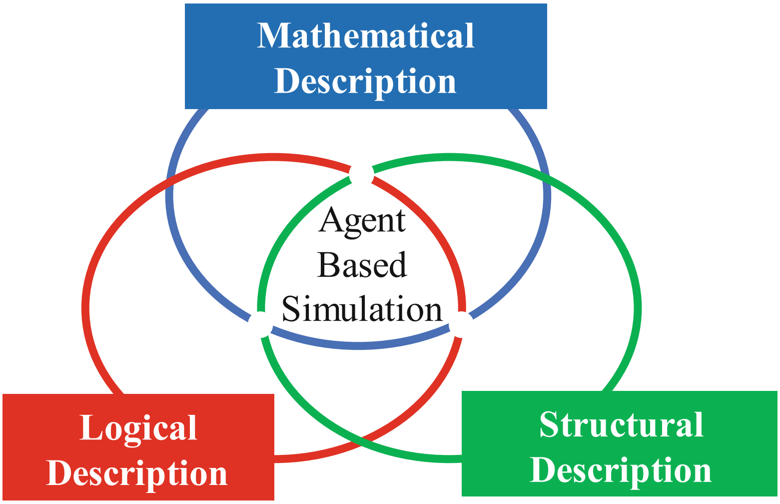

Fig. 6. Systematic ordering of models handled in education

An agent does not belong to one of the three categories. Basically, it has a logical or structural description (or both) and it is re-expressed with a mathematical description so that it can be incorporated into agent behavior. In other words, it is created from a combination of at least two or more descriptions. To apply ABS, the phenomenon must be such that at least a structural or logical description can be made and incorporated into a mathematical description. When Yano et al. modeled the influence between students and teachers, they created a logical description model of the relationship between various elements relating to student learning and the amount of learning. After that, they created a mathematical description model to express it.

As an example of educational policy implementation, on the other hand, results have been obtained indicating that small class sizes contribute to improvement of the students' self-image with respect to their academic learning [34]. Since self-image is difficult to describe by a mathematical model, however, it cannot be incorporated into the process of formulating educational policy using ABS.

6 Conclusion

This paper has discussed the use of ABS in addressing questions about why specific outcomes occurred following policy implementation, and whether a process of formulating education policy can be created to reduce the risks in formulation of policy. Answers were derived from cases in which the ABS approach was realized in ways that matched relatively closely with educational policy.

ABS was positioned as a middle-range model, and certain limitations were noted. In terms of the level of description, the approach is limited to policy addressing matters that are subject, at the very least, to a structural or logical description and that can be rewritten into a mathematical description. In formulating educational policy, however, the data are often insufficient and many challenges arise. In that context, ABS is a powerful tool for the interpretation of scenarios. If an environment was created that supports a tool like this, then educational policy could undergo major changes in the future.

Acknowledgments. We thank all reviewers for their valuable comments, and Michelle Pascoe, PhD, from Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

References

1. Tadashi, S.: Yutorikyoiku no rinen ni kannsuru kousatstu. Kanagawa daigaku•kyouiku kennksyuu **32**, 49–53 (2012). (in Japanese)
2. Sakakibara, E.: “Yotorikyouiku” de nihon metsubou. In: Bungeishunju (eds.) The Issues in Education for you Children. Bungeishunju, Tokyo (2001). (in Japanese)
3. PUBLIC LAW 107–110—JAN. 8, 2002 (2002)
4. Belfield, C.R., Nores, M., Barnett, S., Schweinhart, L.: The high/scope perry preschool program cost–benefit analysis using data from the age-40 followup. J. Hum. Resour. **41**(1), 162–190 (2006)
5. Gilbert, N.: Agent-Based Model, Quantitative Applications in the Social Sciences, vol. 153. SAGE Publications, London (2007)
6. UNESCO Bangkok: UNESCO Handbook on Education Policy Analysis and Programming, Volume 1: Education Policy Analysis. UNESCO Bangkok (2013)
7. Shum, B.S.: UNESCO Policy Brief: Learning Analytics. Published by the UNESCO Instute for Information Technologies in Education, Moscow (2012)
8. Textbook Publishers Association of Japan, Gakusyusidouyouryoutozissizikitokyoukasyonohakkannzikitinokannkeinituitesetumeisitekudasai. http://www.textbook.or.jp/question/answer/a04.pdf. Accessed 20 June 2018. (in Japanese)
9. Ministry of Education, Culture, Sports, Science and Technology, Syougakkou gakusyuu sidou youryo, December 1998. http://www.mext.go.jp/a_menu/shotou/cs/1319941.htm. Accessed 20 June 2018. (in Japanese)
10. Takayama, K.: A nation at risk crosses the Pacific: transnational borrowing of the US crisis discourse in the debate on education reform in Japan. Comp. Educ. Rev. **51**(4), 423–446 (2007)
11. Ministry of Education, Culture, Sports, Science and Technology, Gakusyusidoyoueyoutounokyouikukateinokizyuntounoarikatanitsuite, http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/021/siryo/attach/1399083.htm. Accessed 20 June 2018. (in Japanese)

12. Hanushek, E.A., Kain, J.F., Rivkin, S.G., Branch, G.F.: Charter school quality and parental decision making with school choice. J. Public Econ. **91**(5), 823–848 (2007)
13. Saunders, W.M., Goldenberg, C.N., Gallimore, R.: Increasing achievement by focusing grade-level teams on improving classroom learning: a prospective, quasi-experimental study of Title I schools. Am. Educ. Res. J. **46**(4), 1006–1033 (2009)
14. Schelling, T.C.: Dynamic models of segregation. J. Math. Sociol. **1**(2), 143–186 (1971)
15. Yano, Y., Kanzawa, A., Yamada, T., Yoshikawa, A., Terano, T.: Does small-size class improve academic abilities of students? —An analysis by agent-based approach—. Trans. Jpn. Soc. Inf. Syst. Educ. **32**(4), 236–245 (2015). (in Japanese)
16. No Child Left Behind Act Is Working. https://www2.ed.gov/nclb/overview/importance/nclbworking.pdf. Accessed 20 June 2018
17. Riddle, W.: What Impact Will NCLB Waivers Have on the Consistency, Complexity and Transparency of State Accountability Systems?. Center on Education Policy, Washington, DC (2012)
18. Finnigan, K.S., Gross, B.: Do accountability policy sanctions influence teacher motivation? Lessons from Chicago's low-performing schools. Am. Educ. Res. J. **44**(3), 594–630 (2007)
19. Dillon, S.: Schools Cut Back Subjects to Push Reading and Math. New York Times, New York (2006)
20. Meier, D.: Many Children Left Behind: How the No Child Left Behind Act is Damaging Our Children and Our Schools. Beacon Press, Boston (2004)
21. Sklar, E., Davies, M., Co, M.S.T.: SimEd: simulating education as a multi agent system. In: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 3, pp. 998–1005. IEEE Computer Society, Washington, DC (2004)
22. Word, E., et al.: Student/Teacher Achievement Ratio (STAR) Tennessee's K-3 Class Size Study. Final Summary Report 1985-1990. Tennessee State Department of Education (1990)
23. Konstantopoulos, S.: How consistent are class size effects? Eval. Rev. **35**(1), 71–92 (2011)
24. Molnar, A., Smith, P., Zahorik, J., Palmer, A., Halbach, A., Ehrle, K.: Evaluating the SAGE program: a pilot program in targeted pupil-teacher reduction in Wisconsin. Educ. Eval. Policy Anal. **21**(2), 165–177 (1999)
25. Johnson, R.T., Johnson, D.W.: Cooperative learning in the science classroom. Sci. Child. **24**(2), 31–32 (1986)
26. Aronson, E., Patnoe, S.: Cooperation in the Classroom: The Jigsaw Method, 3rd edn. Pinter & Martin Ltd., London (2011)
27. Kuniyoshi, K., Kurahashi, S.: How do children learn and teach? In-class collaborative teaching simulation on the complex doubly structural network. SICE J. Control. Meas. Syst. Integr. **10**(6), 520–527 (2017)
28. Spoelstra, M., Sklar, E.: Agent-based simulation of group learning. In: Antunes, L., Paolucci, M., Norling, E. (eds.) MABS 2007. LNCS (LNAI), vol. 5003, pp. 69–83. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-70916-9_6
29. Kunigami, M., Kobayashi, M., Yamadera, S., Yamada, T., Terano, T.: A doubly structural network model and analysis on the emergence of money. In: Takadama, K., Cioffi-Revilla, C., Deffuant, G. (eds.) Simulating Interacting Agents and Social Phenomena, Agent-Based Social Systems, vol. 7, pp. 137–149. Springer, Heidelberg (2010). https://doi.org/10.1007/978-4-431-99781-8_10
30. Slavin, R.E.: Student teams and achievement divisions. J. Res. Dev. Educ. **12**(1), 39–49 (1978)
31. Arai, A., Terano, T.: *Yutori* is considered harmful: agent-based analysis for education policy in Japan. In: Shiratori, R., Arai, K., Kato, F. (eds.) Gaming, Simulations, and Society, pp. 129–136. Springer, Tokyo (2005). https://doi.org/10.1007/4-431-26797-2_14

32. Mitchell, D.E., Beach, S.A., Badaral, G.: Modeling the relationship between achievement and class size: a re-analysis of the Tennessee project star data. Peabody J. Educ. **67**, 34–74 (1989)
33. Yoshikawa, A., Takahashi, S.: Review on educational application of agent technologies. Trans. Jpn. Soc. Inf. Syst. Educ. **35**(1), 5–12 (2018). (in Japanese)
34. Mulkey, L.M., Catsambis, S., Steelman, L.C., Crain, R.L.: The long-term effects of ability grouping in mathematics: a national investigation. Soc. Psychol. Educ. **8**(2), 137–177 (2005)