THE UNIVERSITY OF SYDNEY

Automated Essay Scoring in Australian Schools: Collective Policymaking

Policy Brief

November 2022

Sydney Social Sciences and Humanities
Advanced Research Centre

Education Innovations Policy Brief Series
ISSN 2653-6757

Sydney Social Sciences and Humanities Advanced Research Centre (SSSHARC)

Sydney Social Sciences and Humanities Advanced Research Centre is a collaborative research facility designed to advance innovative research partnerships and methodologies across the humanities and social sciences. We have a commitment to research excellence and knowledge sharing. We broker and advance research between researchers at the University of Sydney and their outside interlocuters via visiting fellowships, exchanges, and a range of in-house initiatives designed to build authentic research partnerships that will grow and change to meet the demands of the future.

More information is available at:

sydney.edu.au/arts/our-research/centres-institutes-and-groups/sydney-social-sciences-and-humanities-advanced-research-centre/about-sssharc.html

Education Innovations and Education Futures Studio

Education Innovations is a SSSHARC-funded initiative led by Professor Kalervo Gulson at the University of Sydney. Central to this initiative is the Education Futures Studio (EFS), where diverse stakeholders come together to experiment with contemporary technologies of governance in education. At the EFS, we try to unpack and contextualise socio-technical controversies in education. We advance collaborations across the education sector in order to explore the links between teaching, emerging technologies, policy, learning, and research.

More information is available at:

education-futures-studio.sydney.edu.au

Acknowledgements

This project was funded by the Australian Research Council (Grant ID: FT180100280) and Sydney Social Sciences and Humanities Advanced Research Centre (SSSHARC). We thank everyone who participated in the hybrid forum for sharing their ideas and diverse expertise, which informs this policy brief. We also appreciate the support of the SSSHARC team, the Faculty of Arts and Social Sciences, and the University of Sydney.

Authors

- Professor Kalervo Gulson, Sydney School of Education and Social Work, University of Sydney.
- Professor Greg Thompson, Education Research, Queensland University of Technology.
- Dr Teresa Swist, Senior Research Officer, Education Futures Studio, University of Sydney.
- Associate Professor Kirsty Kitto, Connected Intelligence Centre, University of Technology Sydney
- Professor Leslie Rutkowski, Quantitative Research Methodology, Indiana University.
- Professor David Rutkowski, Educational Research Methodology, Indiana University.
- Dr Anna Hogan, Senior Research Fellow, School of Teacher Education and Leadership, Queensland University of Technology.
- Vincent Zhang, Co-founder of Tech for Social Good, a non-profit for young Australians interested in responsible technology.
- Dr Simon Knight, Senior Lecturer, Director of the Centre for Research on Education in a Digital Society, University of Technology Sydney.

Recommended citation

Gulson, K., Thompson, G., Swist, T., Kitto, K., Rutkowski, L., Rutkowski, D., Hogan, A., Zhang, V., Knight, S. (2022). *Automated Essay Scoring in Australian Schools: Collective Policymaking* . Policy Brief, November 2022. Education Innovations Policy Brief Series ISSN 2653-6757. Sydney Social Sciences and Humanities Advanced Research Centre (SSSHARC), University of Sydney, Australia.

Automated Essay Scoring in Australian Schools: Collective Policymaking

Policy Brief

Contents

Executive summary

This summary outlines critical issues associated with the use of Automated Essay Scoring (AES) technology in the Australian education system. The key insights presented in this paper emerged from a collaborative, multi-stakeholder workshop held in July 2022 that explored an automated essay-scoring trial and generated future possibilities aligned with participant interests and expertise. Drawing on the workshop and our expert understanding of the wider landscape, we propose recommendations that can be adopted by various stakeholders, schools, and educational systems.

There are compelling reasons for Australian schools and education departments to investigate the use of AES. AES could potentially alleviate aspects of teachers' workload at a time when teacher attrition is historically high and teacher recruitment historically low. At the same time, AES also has the potential to de-professionalise and deskill of teachers. Educationalists are acutely aware that quality feedback can help students improve their learning across multiple subjects and domains, however parents and many are reluctant to hand that responsibility over to AES.

In 2018, concerns among teachers, teachers' unions, principles, and parents became apparent when the federal Department of Education, Skills and Employment attempted to implement a form of AES in The National Assessment Program – Literacy and Numeracy (NAPLAN). These concerns primarily registered around three issues:

- de-professionalisation of teachers,
- inequitable infrastructure in Australian schools, and
- lack of transparency from examination authorities as to how marking decisions are made.

AES cannot be approached in one dimension. It is crucial to frame the use of AES in schools as an issue with interrelated ethical, social, technical, and political implications.

The use of AES in NAPLAN ultimately proved to be politically unpopular, leading to its suspension. However, the growing implementation of AES in schools across the globe means that the use of this technology is likely to re-emerge as a controversial issue in Australia. Without political leadership in this area, it is ultimately up to educational institutions and agencies, policymakers, and school communities to assess the benefits and pitfalls of AES and navigate the way forward. Our recommendations will assist the emergence of good governance in this area.

To begin, it is crucially important to identify whether AES will be used in high-stakes or low-stakes tests. High-stakes tests are defined as those with consequential outcomes for students or educators, such as the determination of progression of students or rankings of school institutions.

If AES is to be used in Australian schools, the following issues must be considered:

- the capacity of stakeholders, including principals, teachers, and parents, to understand how AES systems work
- the infrastructure required to support the use of AES
- the potential impacts of AES on assessment and workload practices which requires adequate professional development resources
- competing interests and values between schools, departments, and institutions associated with using AES
- how the use of AES relates to and integrates with broader policy frameworks.

The investigation of these issues requires information sharing, dialogue, and negotiation among diverse stakeholders, including teachers, parents, students, leaders, and policymakers.

In addition to this engagement, schools and other educational institutions must also discuss the implementation of AES tools with AES system developers and commercial vendors, so as to better understand the functions and limitations of the AES tool, as well as its implications for professional and assessment practices. Only then can decision-makers evaluate whether a specific AES system is worth the investment of funds and resources, including teacher workload, in both the medium and longer term.

Although it appears as yet another drag on teacher time, the participatory and collaborative development of AES guidance, policy, and regulation is crucial. It ensures that pluralistic views and shared values are reflected in any innovations or reforms across the education sector. To ensure a collaborative foundation, the introduction of AES must be informed by stakeholder expertise across multiple locations and decision-making levels, including classrooms, schools, organisations, and state, territory, and national jurisdictions. **For Australia, we recommend multi-scalar policy development informed by educators, policymakers, and representatives from educational technology companies engaging in cooperative learning and action.**

When is AES not recommended?
The use of AES is not recommended for high-stakes tests in schools in most instances.

Where high stakes are likely for students, teachers, principals, or school communities, how AES works and the ethical implications of its use must be clearly explained.

Summary infographic

The key issues, takeaways, critical questions, and recommendations are outlined here. Please refer to the related white paper[1] for further details. These insights can support diverse stakeholders to navigate the interrelated social, technical, ethical, and political dimensions of AES systems in high-stakes education contexts.

01

Issue 1: AES system complexity and contexts

Key takeaway

More learning about AES system complexity and contexts is needed.

Critical questions

Has the AES vendor been transparent about any immediate or longer-term social and ethical impacts?

Are you aware of publicly available learning tools that would help people to understand AES systems?

Recommendations

Make learning tools that enable diverse stakeholders to understand how AES systems work.

02

Issue 2: School infrastructure capacity to deploy AES

Key takeaway

The range of infrastructure and capacity-building to fully support AES across Australian schools is underexplored.

Critical questions

Is there adequate resourcing, internet reliability, and labour to maintain the digital infrastructure and introduced system?

Do all people in your school or organisation have access to the required technology and skills to implement AES?

Recommendations

Identify the digital infrastructure and skills required to support the use of AES across urban, regional, and remote schools.

1 Gulson, K., Thompson, G., Swist, T., Kitto, K., Rutkowski, L., Rutkowski, D., Hogan, A., Zhang, V., Knight, S. (2022). *Automated Essay Scoring in Australian Schools: Key Issues and Recommendations*. White Paper, November 2022. Education Innovations White Paper Series ISSN 2653-6749. Sydney Social Sciences and Humanities Advanced Research Centre (SSSHARC), University of Sydney, Australia.

03

Issue 3 : Impact of AES upon professional practice

Key takeaway

Assessment and workload practices are being displaced with AES, so there is a need to expand teachers' socio-technical expertise and to build co-designed systems.

Critical questions

Does your school or organisation provide opportunities to discuss the positive and negative impacts of new technology, such as AES, on professional practices?

Would you attend professional development that provided opportunities to learn about and experiment with automated technologies like AES?

Recommendations

Prioritise professional development and co-designed AES systems which value, and build upon, teachers' judgement and socio-technical expertise.

04

Issue 4 : Cross-sectoral interests and values associated with AES

Key takeaway

Opportunities to explore multiple stakeholder interests and values about AES are currently lacking.

Critical questions

Do you know the details of why, when, and where an AES system is introduced?

Who decides if and how an AES system is introduced into your jurisdiction, organisation or school?

Recommendations

Provide opportunities for sharing knowledge and decision-making about the use of AES between diverse stakeholders.

05

Issue 5 : Policy uncertainty regarding AES and emerging EdTech

Key takeaway

Existing policies are not keeping pace with rapid technological change, such as AES, in Australian schools.

Critical questions

Do you know what policies frame the introduction and use of AES or other education technologies in your jurisdiction, organisation, or school?

What avenues are there for appeals, or new approaches, to be made about AES decisions and systems?

Recommendations

Connect and integrate policies for the use of AES in high-stakes education contexts.

Automated Essay Scoring (AES) sphere of impact

Recognising that the AES sphere of impact (Diagram 1) is distributed across multiple locations and jurisdictions is crucial to collective policy development and implementation in this area.

Classrooms/Schools

Most AES systems are commercial education technology packages purchased by schools who receive no guidance about product quality or the benefits or drawbacks of using the technology in specific school settings. Procurement processes and guidelines must be developed with input from teachers and the broader school community to help evaluate the benefits, and potential harms, of introducing automated technologies that consume limited resource and place new demands on teachers and administrators.

Regional/State/Territory

The speed of science, technology and innovation development poses challenges for education departments, including training, quality assurance and evaluation. These challenges are exacerbated by digital infrastructure gaps and diverse needs, including for rural and regional areas, and schools with high proportions of low socio-economic background students. To ensure digital inclusion across Australian schools, infrastructure and capacity-building needs across different locations must be closely mapped. Successful AES implementation requires both technical and social capabilities.

National/Federal

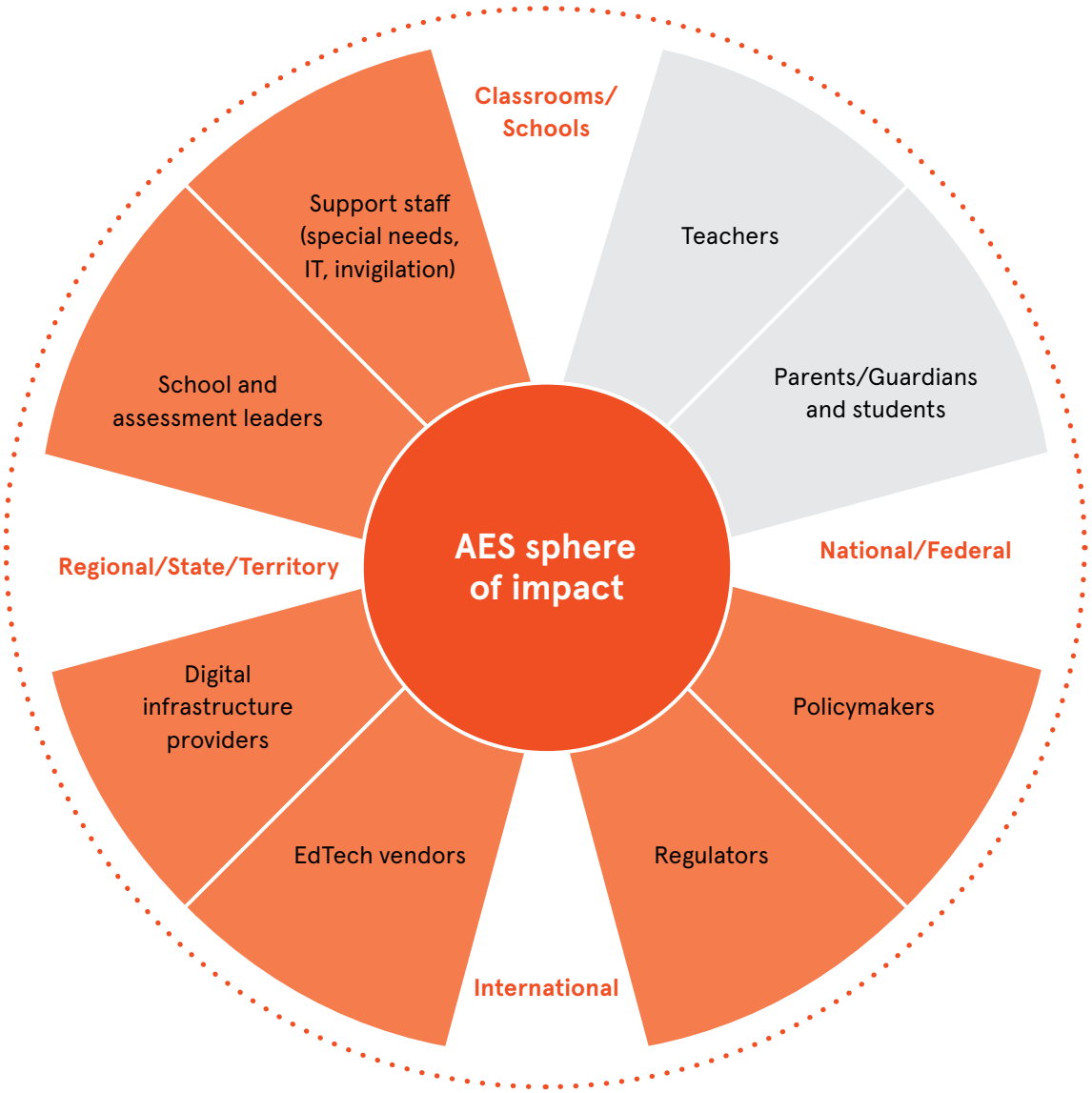
The introduction of AES in schools requires broader policy changes around assessment and the use of artificial intelligence (AI) in schools. In Australia, existing policies relating to the use of AI in education are significantly underdeveloped compared to other countries and regions. Current National School Reform Agreement initiatives will expand digitalisation and interoperability across national assessment, tracking, and data assets[2]. Where AES fits into these policy priorities remains to be seen especially in relation to acceptable and non-acceptable levels of automation across school infrastructure and assessment processes.

International/Global

The use of AES in education settings is currently most widespread in the United States where it is deployed in a range of high-stakes contexts that carry consequential outcomes for students and schools. In the European Union, AES in education is seen as a high-risk use of artificial intelligence. The lack of coordinated global policy around AI in education is beginning to be addressed by international and regional organisations such as UNESCO and the Organisation for Economic Cooperation and Development. Australia could also learn valuable lessons from examining large-scale AES rollouts in the US and other jurisdictions. This would increase our capacity to assess the ethical, legal, and financial implications of such rollouts and their impact on school, departmental, and regional decision-making.

2 Australian Government (2022). The National School Reform Agreement. Department of Education. https://www.education.gov.au/quality-schools-package/national-school-reform-agreement

Diagram 1: AES sphere of impact



Collective policymaking

The collective design and development of AES policy and governance can be applied in the Australian education context in multiple ways (Diagram 2).

Implement participatory procurement processes

Product procurement in the education sector is currently steered by EdTech company priorities. A more accountable and transparent agreement needs to be developed across the sector to establish minimum standard and benchmarking[3]. As more AES technologies are brought to market, there is a need for ethical procurement guidelines to be developed with diverse education stakeholders. These guidelines must include responsibilities of companies explain what the technology does in the educational context in which it will be applied, including specific schools. This process should begin with a scoping of procurement policies currently in place in departments, and any issues that have already arisen in relation to proprietary technology.

Build upon best practice guidelines

Transparency in process is essential moving forward. In the US, standards of best practice in automated scoring have been developed[4]. To align with established best practice, the use of AES – whether in high-stakes or low-stakes contexts – requires full technical documentation to be made available and evidence to validate all claims made for the technology.

Create education-specific, risk-based frameworks

The European Union has developed a risk-based framework for evaluating the use of AI in different social policy areas. This framework is designed to foster public trust in AI systems based on four categories: unacceptable risk, high risk, limited risk, and minimal risk[5]. The EU framework identifies the use of AI in education as a high-risk activity if used in ways that "may determine the educational and professional course of a person's life"[6]. The creation of education-specific frameworks to evaluate the risk associated with the implementation of emerging technologies, such as AES systems, in different Australian education settings is imperative.

Develop AES-focused audits

The UK has advised that algorithmic audits become commonplace especially in high-stakes areas like education. 'Algorithmic auditing refers to a range of approaches to review algorithmic processing systems. It can take different forms, from checking governance documentation, to testing an algorithm's outputs, to inspecting its inner workings'[7]. While there is a suggestion that these audits can be part of a regulatory environment, undertaking audits is dependent on having technical expertise that is unlikely to exist in a school or even an education system. It is therefore allowable that the review of the use of AI systems in high stakes areas can be undertaken by a third-party audit.

3 Hillman, V. (2022). EdTech procurement matters: It needs a coherent collusion, clear governance and market standards. LSE Department of Social Policy, Working Paper April 2022.https://www.lse.ac.uk/social-policy/Assets/Documents/PDF/working-paper-series/02-22-Hillman.pdf

4 Wood, S., Yao, E., Haisfield, L., and Lottridge, S. (2021). Establishing Standards of Best Practice in Automated Scoring. ACT Research, Technical Brief, July. https://www.act.org/content/dam/act/unsecured/documents/R2100-auto-scoring-standards-2021-07.pdf

5 European Commission (2021). Proposal for a Regulation laying down harmonised rules on artificial intelligence. https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence

6 Ibid, p.26

7 UK Government (2022). Auditing algorithms: the existing landscape, role of regulators and future outlook. https://www.gov.uk/government/publications/findings-from-the-drcf-algorithmic-processing-workstream-spring-2022/auditing-algorithms-the-existing-landscape-role-of-regulators-and-future-outlook

Prototype policy ideas relating to AES

Policy prototyping is a process of experimenting and testing policy ideas to improve existing governance mechanisms [8]. More recently, policy prototyping has been used by transnational technology corporations, such as Meta (formerly Facebook) to establish rule and law-making processes in partnership with governments, academia, and civil society[9]. This process could potentially be adapted to the education sector as a way for diverse stakeholders to collectively generate ideas which address the range of ethical, social, and political issues associated with AES systems.

Initiate an independent advisory body for large-scale assessment in Australia

There is currently no independent governing body in Australia for large-scale national assessment. In the US, the National Assessment of Educational Progress (NAEP) is the largest body for continuing assessment. It is "an independent, bipartisan organisation made up of governors, state school superintendents, teachers, researchers, and representatives of the general public" which sets NEAP policy[10]. It is suggested that an independent advisory body for large-scale assessment in Australia be initiated along similar lines. This group would have oversight for key developments in large-scale assessment and reporting in Australia.

Establish a collective evidence platform that shares systemic information about Australia's education sector

There is currently no knowledge-sharing platform that communicates systemic information about Australia's education sector. In the US, the National Centre for Education Statistics provides evidenced-based reports on a range of indicator topics around the 'condition of education', such as district-level, socio-economic differences, internet access, teacher/staff characteristics and turnover, student learning, high school completion, revenue sources and expenditure [11]. A similar knowledge-sharing platform in the Australian context would allow the sector to understand and respond to the many place-based, infrastructural, cultural and finance-related factors that influence education and assessment outcomes in Australia, including the use of AES.

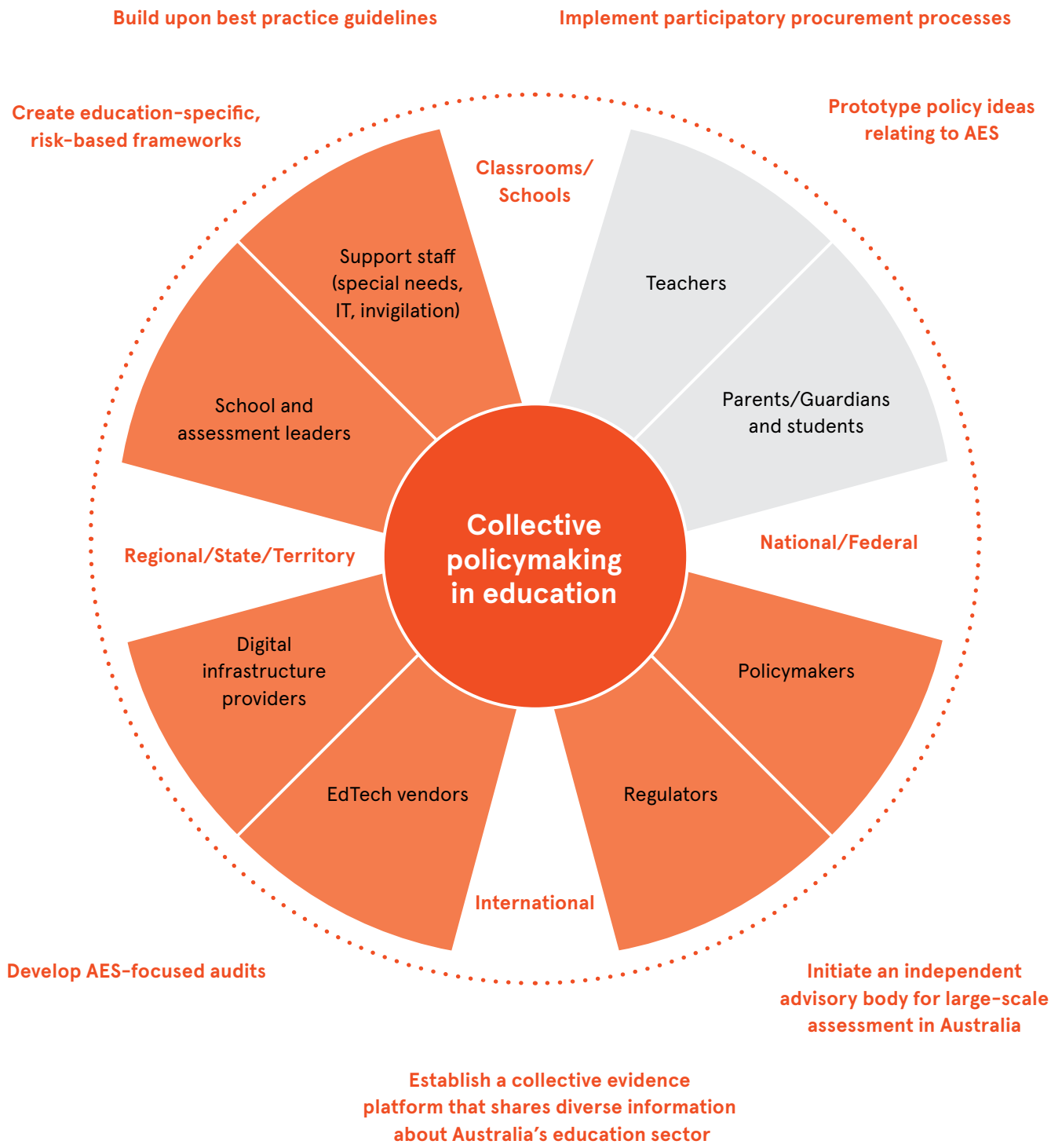
8 Stanford Law School (2018). Prototyping in Policy: What For?! https://conferences.law.stanford.edu/prototyping-for-policy/2018/10/22/prototyping-in-policy-what-for/

9 Open Loop (2021). Let's Experiment. https://openloop.org/lets-experiment/

10 National Center for Education Statistics. (NCES 2022). https://nces.ed.gov/nationsreportcard/ National Assessment of Education Progress.

11 Ibid.

Diagram 2: Collective policymaking in education



Conclusion

The complexity of AES systems cannot be resolved by a single stakeholder or siloed solutions. Collective policymaking across scales is necessary to identify networked tensions and possibilities in education.

Interconnected learning, experimentation, and policymaking is urgently required. Multidimensional and multi-scalar action must happen now.

We encourage readers to consider how they can identify opportunities to propose, or implement, collective policymaking processes. Next actions are:

- Share this policy brief with your networks who would like to trial and test collective policymaking.
- Connect with leaders who can allocate funding and resources to implement ideas.
- Communicate your collective policymaking experience (successes and failures), from which others can learn.
- Build stakeholder networks across policymaking scales which can inform collective learning, experimentation, and action.

We hope you are inspired to explore collective policymaking so that present and future uses of AES systems in Australia can serve the interests of diverse stakeholders.

AES cannot be approached in one dimension. It is crucial to frame the use of AES in schools as an issue with interrelated ethical, social, technical, and political implications.

THE UNIVERSITY OF SYDNEY

Sydney Social Sciences and Humanities Advanced Research Centre

:Education Futures Studio]

More information
If you would like further information about our Education Innovations Policy Brief Series or the Education Futures Studio, please contact:

Kalervo Gulson
kalervo.gulson@sydney.edu.au

Greg Thompson
g6.thompson@qut.edu.au

Teresa Swist
teresa.swist@sydney.edu.au

We acknowledge the tradition of custodianship and law of the Country on which the University of Sydney campuses stand. We pay our respects to those who have cared and continue to care for Country.

CRICOS 00026A