# Scaling up Women in Computing Initiatives: What Can We Learn from a Public Policy Perspective?

Elizabeth Patitsas, Michelle Craig, and Steve Easterbrook
Department of Computer Science
University of Toronto
Toronto, Ontario, Canada
patitsas, mcraig, sme@cs.toronto.edu

## ABSTRACT

How to increase diversity in computer science is an important open question in CS education. A number of best practices have been suggested based on case studies; however, for scaling these efforts up in a sustainable fashion, it remains unclear which types of initiatives are most effective in which contexts. We examine gender diversity initiatives in CS education from a policy analysis perspective, adapting McDonnell and Elmore's 1987 notion of policy instruments, wherein the initiative is the unit of analysis. We present a conceptual framework for categorizing the different policy instruments by a cross of 'leverage' and 'targetedness', and discuss how different types of initiatives will scale. We argue that universally-targeted, high-leverage initiatives are most important for scaling up diversity initiatives in CS education, with medium-leverage change being a stepping stone to high leverage change.

## 1. INTRODUCTION

In the past three decades, a great deal of effort has been put into trying to improve female participation in computer science. Yet, the numbers in North America haven't budged: women continue to make up only 18% of CS majors [21].

Some efforts have had tremendous, sustained results. For example, Carnegie Mellon and Harvey Mudd have both increased the percentage of women studying CS from around 15% to around 40% in the span of a few years [1, 38].

These initiatives remain unusual, however. While they provide proof that change can happen, they do not provide a roadmap on how to bring that change to scale.

Scale has become a new focus for CS education [28, 20]: as CS is increasingly taught to a wider audience – especially in k-12 school systems – how can we handle the scale? To look at the issue of scale, we adopt the lens of public policy analysis: we consider women-in-computing initiatives as acts of policy.

Researchers who study education at scale – particularly education policy – often work with units of analysis such as initiatives, policies, schools, or regions. In comparison, in the CS education community we tend to work with individuals as our units of analysis. Even when we evaluate initiatives, we tend to evaluate one initiative at a time, using the beneficiaries of the initiative as the unit of analysis.

In this paper we will be considering initiatives as the unit of analysis, rather than individuals. We present a conceptual framework for classifying diversity initiatives, providing a first step toward a policy analysis approach to computer science education.


*ICER '15,* August 09 - 13, 2015, Omaha, NE, USA.

ACM 978-1-4503-3630-7/15/08 ...$15.00
http://dx.doi.org/10.1145/2787622.2787725.

## 2. POLICY INSTRUMENTS

In this paper we treat women-in-computing initiatives as acts of department or classroom policy. This framing of diversity initiatives allows us to draw on literature from public policy analysis. To simplify the scope of this paper, when we talk about 'diversity' we will focus on gender diversity specifically. We will return to the broader diversity issues in subsection 7.1.

We use a broad definition of 'initiative': any formal effort to increase female engagement in computing. Some examples of what we mean by initiative, or policy, include:

**Admissions criteria change:** changing admissions criteria to focus on 'non-numerics' like at CMU [38]

**Degree requirement change:** having multiple CS1s separated by experience level / applications [1]

**Curriculum change:** using MediaComputation to teach CS1 in a context-focused fashion [26]

**Pedagogy change:** randomly calling on students to ensure that all students speak equally in the classroom and overcome a 'defensive climate' [24]

**Sending students to the Grace Hopper Celebration:** to foster community amongst female students and expose them to the 'real world' of computer science [1]

**Research opportunities for first-years:** to foster early interest in CS [1]

**K-12 outreach:** bringing k-12 students to the university to expose them to computer science activities, such as via a summer camp or day-camp [14]

The education policy literature provides us a notion of 'policy instrument': thinking about qualities of policies themselves, using the policies as units of analysis. The approach comes from McDonnell and Elmore's 1987 classification of macro-level policies as being mandates, initiatives, system change, or capacity-building [40].

Other policy researchers have classified policies differently (e.g. [18]); the insight here is that policies themselves can be classified and their classifications theorized upon.

As our focus here is on department-level policies – rather than nation/state-level – with a focus on scaling up women-in-computing initiatives, we have constructed our own conceptual framework of policy instruments. We classify women-in-computing initiatives by two axes: 'targetedness' (how broad the audience is) and 'leverage' (how deeply the system is changed).

# 3. TARGETEDNESS

In public health, the Universal/Selective/Indicated (USI) model has proven to be an effective conceptual tool for forming public health initiatives [51]. In this model, initiatives are categorized by the intended audience: universal strategies are aimed at whole populations; selective strategies are aimed toward at-risk groups; and indicated strategies are aimed toward individuals displaying signs of the condition in question.

To give some more concrete examples from suicide prevention:

**Universal initiatives:** restricting exposure to suicide content in mass media, adding barriers on bridges, and restrictions on pharmaceutical dispensing [17]

**Selective initiatives:** selective initiatives here include suicide prevention centres and hot-lines, community or school suicide prevention programmes, and programmes for veterans and military personnel [17]

**Indicated initiatives:** training general practitioners to spot warning signs in patients and how to talk to patients about suicide, postvention, and crisis hot-lines [17]

Here we'll extend the USI model from public health to an education setting. We will refer to the spectrum it represents as 'targetedness': with universal initiatives being less 'targeted' and indicated ones being most 'targeted'. 'Targetedness' refers to how *wide* the audience is, not which audience is being targeted.

## 3.1 Universal initiatives

In an education context, our idea of "population" differs. A "population" can be a whole classroom of students, a whole CS department or school – or even the general population of a country. The key notion is that universal initiatives are carried out without regard to population target groups or risk factors. Examples of diversity initiatives in CS education which are universal include:

- A CS department makes a mentorship programme available to all students
- Pair-programming and peer instruction for a whole classroom
- A university mandates that all students need to take CS, and its CS department provides multiple, engaging, versions of CS1 that are tailored to different students' interests – like at HMC
- A department changes their admission process that affects all CS students – like at CMU
- A conference switches to using blind review of its submissions.
- A CS professor implements a 'social-psychological intervention' in their classroom (e.g. values-affirming essay), to improve the self-efficacy of all students [53]

Each of these initiatives affects a differently sized population, but the initiative affects all members of its population. It is worth noting that all of the above practices are known (or thought to) to improve female representation in CS. They disproportionately benefit women and other minorities, but also aid majority-members. The same is true of the universal initiatives for suicide prevention: restricting suicide content in the mass media affects all mass media consumers, but disproportionately helps those with suicidal ideations.

While it may seem quite costly to run a universal initiative, given that it has to reach the whole population, recent meta-reviews in public health have found universal initiatives are actually the most cost-effective: "a large number of people at small risk may give rise to more cases of disease than a small number who are at high risk" [17]. And as many diseases are contagious, universal initiatives can improve the resilience of the whole population.

Furthermore, as universal initiatives target the whole population, they provide a means of reaching at-risk individuals who are not in contact with institutional services [17].

## 3.2 Selective initiatives

In comparison, selective initiatives target a population known to be underrepresented in CS; they specifically and explicitly benefit that group, and provide them with targeted support to 'level the playing field' with dominant groups in CS. Examples include:

- A CS department makes a mentorship programme available to all female students
- Departmental women-in-CS clubs
- Giving the opportunity for female students to go to the Grace Hopper Celebration
- Outreach initiatives for school-age girls
- Scholarships for women in CS

Many selective initiatives in public health – suicide-related or otherwise – have been found effective. Meta-reviews have noted that long-term selective initiatives tend to be more successful than short-term ones. Selective initiatives need to be culturally and contextually appropriate to the audience(s) in order to be effective [17].

Certain selective initiatives have been noted as having potential harm – for example, being seen associated with a group for a stigmatized disease/condition. Like in public health, stigmatization has the potential to be an issue in education. Audit studies have found that job candidates associated with affirmative action (which targets specific groups) are perceived as less competent than identical job candidates without those associations [31]. This effect is strongest when the job candidate's competence is ambiguous [31]. Some qualitative studies of women in science have also noted that beneficiaries of research grants for women in physics feel they are perceived as less competent for receiving the "women's award" rather than a traditional research grant [48].

## 3.3 Indicated initiatives

Finally, some examples of indicated initiatives in CS education would include:

- A CS department makes a mentorship programme available to students who have been flagged as struggling in their studies
- A teacher takes the time to encourage a student to study more CS
- An academic adviser notices a student is lacking motivation to study CS, and takes the student to Grace Hopper with them
- A supervisor notices a student is facing sexual harassment in their research lab, and makes appropriate steps to protect the student

The effect of indicated initiatives can be quite strong for the individuals it affects: one-on-one encouragement is a strong indicator of whether black students will take CS [42, 55, 28]; it is strongly beneficial to women also [42].

Indicated initiatives, however, rely on educators to be able to recognize students who need help, and be able to effectively help them. For us to rely on indicated initiatives requires all (or nearly all) CS educators to take part – and as a result scales poorly.

# 4. LEVERAGE

In contemplating scaling up changes, it is also worth considering whether the changes are system-changing or are relatively superficial. Systems thinking offers the notion of *leverage points*: places and ways one can change a system. Donella Meadows constructed a categorization scheme of types of leverage points, and organized them by how much leverage they have in a system. The list in this section goes from least leverage to most – in other words, how deeply (and effectively) the system is changed.

Jay Forrester, a pioneer of systems thinking, noted that although people in a system often know intuitively where to find leverage points, "more often than not they push the change in the wrong direction" [41]. For example, in one of Forrester's studies of urban dynamics from the 1960s, he found that subsidized low-income housing is a leverage point. However, Forrester's model counterintuitively found that the *less* low-income housing there was in a city, the better off the city was – including the low-income citizens [22]. Many more examples of unintuitive leverage-points can be found in [41].

Thinking about leverage points gives us a tool for identifying when changes could be superficial – so that we can focus resources on deeper changes. It also allows to better understand and describe system changes.

It's worth noting upfront that high-leverage changes are often the hardest to make. Systems are resilient and can resist the change; too much change too suddenly can quickly be undone. On the other end, the categories with least leverage – "constants, parameters, numbers" and "the sizes of buffers" – are often superficial. While easier to alter, they rarely lead to systemic change.

## 4.1 Meadows' Leverage Points

**Constants, parameters, numbers:** In systems thinking, systems are thought of as having *stocks* (quantities of things) and *flows* (the altering of stocks). A simple example is a bathtub: there is a flow of water from the faucet into the bathtub, a stock of water in the bathtub, and a flow of water out the drain. Changing the rate of flow in and out of the bathtub has an effect on the system – but does not change the fundamental structure of the system.

Meadows notes that much of politics focuses on this leverage point: how much we spent on $x$, the value of the minimum wage, the value of a tax rate, etc. However, changing the parameter rarely changes the behaviour of the entire system [41]. At the same time, humans tend to focus on parameters [41]; they are concrete and easy to identify.

**Sizes of buffers:** Some systems have *buffers*: stabilizing stocks that are large relative to their flows. Buffers play an important role in many systems – for example, stores keep inventory rather than ordering new stock every time a customer buys something new. The inventory gives the store a buffer from any delays in deliveries or sudden increases in sales.

**The structure of stocks and flows:** The structure of how stocks and flows can have an enormous impact on a system [41]. Redesigning a plumbing system, or refactoring a code base, can have large effects.

**The delays in the system:** Delays in a system affect feedback loops, and can cause unpredictable behaviour in a system. Reducing or increasing delays in the system can have large effects. Often, delays cannot be changed: it takes a fixed amount of time for a baby to mature or for electrons to travel a given distance.

**The strength of negative feedback loops:** A thermostat is a classic example of a system controlled by a negative feedback loop: if it gets too cold, the furnace turns on. If it gets too hot, the furnace turns off. The result is a room with a temperature which varies slightly around a set equilibrium; any disturbance in the equilibrium and it is programmed to return to that state.

The strength of a negative feedback loop is important relative to the impact it is designed to correct [41]. A thermostat may work well on a cold winter day – until somebody opens a window, decreasing the strength of the negative feedback loop.

**The gain of positive feedback loops:** Positive feedback loops are self-reinforcing – such as how the more people catch the flu, the more it spreads; or how the more money you have in the bonk, the more interest you earn. Positive feedback loops can also be known as 'success to the successful': for example, the more research grants a professor receives, the easier it is for them to receive subsequent grants.

**The structure of information flows:** A famous case in energy usage behaviour comes from the Netherlands in the 70s: in a particular suburb of Amsterdam, some otherwise identical houses were built with their electric meters in the basement, and some in the front hall. The houses with the meter in the front hall used one third the electricity as the houses with the meter in the basement, where people rarely saw it. Those who saw their meter every day were hence more conscious of their electricity usage – and used less [41].

Adding a flow of information to the system adds a new loop to the system: it is not increasing the strength of an existing one or its parameters [41]. Removing an information flow is a similarly high leverage change: censorship can have drastic impacts on social systems.

**The rules of the system:** Even more fundamental than the stocks and flows in a system are the rules which govern it: incentives, punishments, constraints, laws, etc. The rules of a system determine its scope and boundaries. Changing a constitution of a country or an organization is an example of this leverage point.

**The power of self-organization:** The system's ability to change its own rules and structures is known as *self-organization*. In a biological context, evolution is an example of self-organization. In a political context, social movements provide a different example.

**The goals of the system:** One of the most fundamental things about a system is its purpose: a school has a goal of teaching students, a hospital of healing patients, a corporation of making profits. If a university changes its purpose from teaching students to producing research – or to making profits – then lower leverage points will be influenced towards that goal.

**The paradigm of the system:** The shared ideas of those in a system – the great unstated assumptions – make up that system's paradigm[1]. Goals are articulated and made within paradigms. People involved in self-organization act in ways affected by their paradigms. And so, the deepest way to change a system is to change the paradigms affecting or defining it.

[1]Meadows lists a final leverage point, 'transcending paradigms', which contradicts the notion of a paradigm, and has been omitted as a result. Her argument that we should 'transcend paradigms' in favour of systems thinking is in itself a reflection of her own paradigm.

## 4.2 Simplifying Meadows: 4 Leverage Groups

To simplify Meadows' leverage-point continuum we group her leverage points into four categories, intentionally borrowing the names from Structure-Behaviour-Function Theory [32]:

**Structural change:** the constants and parameters, the sizes of the buffers, the structure of the stocks and flows.

**System behaviour change:** changing the gains and delays of the feedback loops.

**Function change:** changing how a system is controlled (information flows, rules, self-organization, goals)

**Paradigm change:** changing the very paradigms (in a Kuhnian sense) upon which the system's control is based.

For women-in-computing initiatives, structural changes would include:

- Having CS1 taught by a woman [5] (a parameter change)
- Using female pronouns in assignment instructions [25] (parameter change)
- The size of the departmental Women in CS support group [38] (buffers)
- Assigning groups based on gender [19] (buffers)
- Provide multiple entry points into a CS major [14] (structure of stocks/flows)
- Build "breaks" into the CS1 curriculum as reported in [1] (structure of stocks/flows)

While all these structural changes surely help, they *on their own* do not make for systemic change in CS programmes and classrooms.

System behaviour changes have higher leverage. When it comes to improving diversity in CS, reducing the effects of "success to the successful" makes a major difference. Some system behaviour changes would include:

- Change when students have access to research opportunities [1] (changing delays)
- In CS1, have students write meaningful programs from day one [1] (changing delays)
- Use blind review for scholarship applications [10] (strength of negative feedback loops)
- Reduce and remove potential triggers of stereotype threat, such as posters of Star Trek [13] (strength of negative feedback loops)
- Provide community service learning and co-op opportunities to undergraduates [14] (gain of positive feedback loops)
- Provide more individual encouragement and mentorship to students [14] (gain of positive feedback loops)

Changing how the system is controlled ("function") goes even further in terms of leverage; some examples include:

- Outreach efforts designed to increase/add information flows [14] (information flows)
- More feedback for students (information flows)
- Change entry requirements to the CS major to focus more on 'non-numerics' rather than prior experience [38] (rules)
- Establish a new classroom rule to call on all students randomly, to overcome a 'defensive climate' [24] (rules)
- Perform action research with women and underrepresented minorities in your department [38] (self-organization)
- Empower students to direct some or all of the course content, or use open-ended projects [8] (self-organization)
- Change the goal of the programme to provide an inclusive, positive learning environment for all students [8] (goals)
- Change the learning goals of the class to focus on problem-solving and applications (e.g. MediaComputation [26]) (goals)

And the final leverage point would be paradigm change – some relevant ones would be:

- Shift in thinking: it's the institution that has the problems, not the minority groups [48] (paradigm)
- Shift to an approach to teaching which empowers students, rather than the 'banking model' of education when we deposit 'coins' of knowledge into our students 'bank accounts' [23] (paradigm)
- Shift in thinking: seeing the excellence in computing as something which can be taught/learnt rather than seeing excellence in computing as tied to innate ability (or 'geek genes') [27, 35] (paradigm)

Paradigm changes are difficult to carry out, given the broad change needed to accompany them. When it comes to making change in a system, jumping straight to a paradigm change is usually impractical.

# 5. TARGETEDNESS AND LEVERAGE

The 'targetedness' and 'leverage' qualities of a diversity initiative are independent; in Table 1 we show some examples of initiatives with varying levels of targetedness/leverage.

For example, the table shows different ways an instructor can provide an indicated intervention with a student: giving a student a buffer from a hostile culture is lower leverage (structure); encouragement affects feedback loops and hence provides more leverage (behaviour); providing information on different study and career goals has more leverage (function); and changing their mindset (paradigm) would be a high leverage change.

A particular leverage point can also vary by targetedness: a CS department launching a mentorship programme could open it to all students (universal), open it only to female students (selective), or have it private to students who have been flagged by faculty as needing extra help (indicated).

On the ground, educators spend a great deal of time on indicated initiatives: working with individual students, wondering what to say to them and how to nurture positive growth. Psychology papers often enjoy a lot of attention amongst educators: they focus on these individual changes. As CS educators, when it comes to diversity, we like to talk about issues such as mindset [47], identity [44], stereotype threat [13] and self-efficacy [7]. When we talk about groups, we still talk about the individuals; e.g. "women are more likely to have low-self efficacy". Although these discuss groups, the unit of analysis remains the individual.

While indicated initiatives call for a background in psychology, universal initiatives more often call for a background in sociology. Effective universal initiatives call for thinking about the 'population' in question as a whole, rather than a collection of individuals.

The CS education literature has an understandable tendency to draw upon more disciplinary approaches from psychology than sociology. In Malmi et al's recent survey of theoretical bases of CS education literature, sociology was not even common enough to warrant a category in their data [36]. Often when we talk about universal/selective initiatives we still do so in the language of psychology – increases to self-efficacy of the individuals in a population [7], reducing stereotype threat of the individuals in a group [13], etc.

| | *Universal* | *Selective* | *Indicated* |
|---|---|---|---|
| *Structure* | Whether to use female pronouns in assignment instructions | The size of a women-in-CS club | Add a female student you know is struggling to the women-in-CS club's mailing list |
| *Behaviour* | A CS department provides a mentorship programme for all CS undergraduates | A women-in-CS club provides a mentorship programme | Provide specific encouragement to a student you know has been discouraged by their peers |
| *Function* | Change the learning goals of introductory CS to focus on problem-solving and applications | Change your department's decision making process so a women-in-CS committee provides input on department policies | Counsel a student you know is not engaged with CS about new career goals they can have within the field |
| *Paradigm* | Change the goal of your CS major to promote a collaborative, participatory learning environment | Change the paradigm of a Women in CS club to be intersectional and trans-inclusive | Change a struggling student's mindset from a fixed one to a growth mindset |

**Table 1: Examples of how different initiatives can vary by both targetedness and leverage.**

When we look at cases like HMC [1] and CMU [38], they describe their successes as a series of smaller initiatives, typically with medium or low leverage. To a systems thinker, those smaller initiatives are secondary to the high-leverage changes in organizational goals/paradigms. HMC and CMU both made holistic changes to their CS programmes with the goal of increasing diversity, and making this part of how they teach CS. While not acknowledged in their work, the interaction effects of all their changes is likely greater than the sum of each individual initiative: the changes reinforce each other and change the cultures at those institutions.

# 6. SCALING UP

## 6.1 Targetedness and Scale

When it comes to scaling up, indicated initiatives do poorly: they require nearly every educator to be ready to help a struggling student one-on-one. This requires both buy-in from educators *and* a time commitment from them: some educators may want to help struggling students but not feel they have the resources to do so. Indicated initiatives also suffer from variability: different educators will vary in their ability to diagnose and help different students.

Selective initiatives at first glance look promising for scaling up. Selective initiatives often seem like an obvious choice and follow a clear logic: a group (such as women) is not studying CS, so we should help them. There is a directness to selective initiatives, and it looks good (optics) to those running it.

### 6.1.1 Optics

Selective initiatives win when it comes to optics: a CS department can tout their 'commitment to diversity' by showing off their selective initiatives. Selective initiatives often *look* more like an intended group is being helped than universal initiatives – and meanwhile indicated initiatives are usually invisible to the public.

The optics of selective initiatives can be both a blessing and a curse. The upside is that it can be easier to rally resources and political support to help a disadvantaged group directly – indeed, it's often easier to do so than to change the whole system around their needs. 'Band-aid' solutions are common in policy for good reason: everybody wants to help, but only so much. The downside is that ineffective selective initiatives can act as "pink-washing": superficial efforts used to make people/organizations look good, in turn draining resources from other initiatives and impeding higher leverage change.

### 6.1.2 Illusion of Fairness

Companies which are described as having selective and indicated initiatives are perceived by the public as being fairer companies and better places to work [33].

Problematically, these initiatives can cause an 'illusion of fairness'. In six studies by Kaiser et al [33], participants were grouped in a 2x2 design. Half of the participants were shown information on a fictional company described as having some selective initiative; the other half were given information on a company without any mention of diversity. Then half of the participants in each group were shown evidence that the company they had seen was discriminating against some group (women, blacks); the other half were shown evidence that the company was not discriminating. And then all participants were shown an article about a woman or black man who was suing the company for discrimination.

All participants were then asked to evaluate the company with regard to qualities such as procedural fairness. Troublingly, participants who saw that the company had some diversity initiative thought the company was more procedurally fair, regardless of whether they saw a report showing evidence of discrimination at the company.

Participants were less likely to believe the credibility of the discrimination lawsuit if they saw that the company had some selective initiative. And supporting the paper's findings are a number of legal cases that Kaiser et al reported upon: judges in the US deferring to companies in discrimination cases because the company had enacted selective initiatives – *regardless of how effective they were!* [33].

This phenomenon seems most problematic when the selective initiative is low-leverage. In that case, the lack of real change to the system means that women (or other groups) are still going to be discriminated against — but now they may encounter even more bias *because* of the presence of the selective initiative.

### 6.1.3 Stigma and Stereotypes

Like in public health, there is potential for selective initiatives in CS education to have counter-productive effects for the beneficiaries. Stigma has been reported surrounding receiving women-in-CS 'help': Margolis and Fisher noted a "you're only here because you're a girl" phenomenon [38]. This has been observed in other fields of science: female physics professors who received research grants intended for women found they were taken less seriously as a result and felt a loss of self-efficacy for "needing to get the women's award" rather than a "normal" research grant [48].

Audit studies have found that resumes of women who benefited from initiatives specifically for women are ranked as

less competent than identical resumes without the selective initiatives mentioned [31].

Informing women that they have been selective for special positions to do mathematical work *because* they were women was in one psychology study found to trigger stereotype threat [9] – however, if women were told they were selective based on *both* ability and gender, stereotype threat was not triggered [9].

Yet another social psychological phenomenon associated with selective initiatives is 'subtyping'. When women in CS (or another field) are consistently subtyped as "female computer scientists" rather than "computer scientists", rather than change peoples' ideations of what a computer scientist is, it instead creates a new type in peoples' minds: the *female* computer scientist [45, 6]. This subtype, the *female computer scientist*, is not only separate from the notion of a *computer scientist*, but reinforces that a regular *computer scientist* is not female – effectively further masculinizing the stereotype of computer scientists.

The subtyping effect becomes stronger with repeated exposure. The more effort we put into "women-in-CS" efforts, the more we highlight female computer scientists as *female computer scientists* (rather than computer scientists like any other), the stronger the effect. In short: this effect becomes *worse* with scale.

This is not to say that selective initiatives do not have merit. Instead, organizers of these initiatives need to be cautious of their implementation and wary of the potential side-effects. If selective initiatives are to be scaled up, then even more organizers must be counted on to navigate the potential side-effects.

Universal initiatives do much better with regard to the social psychology around them. If *everybody* receives the same initiative, then you're not making one particular minority group's status salient. Issues of stereotype threat, subtyping and stigma disappear.

#### 6.1.4 Who is Selective?

By helping everybody, universal initiatives also avoid the issue of defining who and who does not fall into a selective group. Feminist theorists such as Judith Butler have well established that gender and sex are both social constructs [11]. If you only offer a programme to 'women', then you need to consider who is a 'woman'. Do you include male-to-female transgender individuals? Female-to-male? When in a trans person's transition do they count (or not count) as a 'woman'? What about intersex individuals (those born with biological aspects of both sexes) or individuals who do not have XX or XY chromosomes?

Other underrepresented minorities in CS suffer from similarly ambiguous boundary lines: race is another social construct with poorly defined boundaries [29]. If you offer a programme for black youth, are half-black youth allowed to attend? One quarter? One sixteenth? What about youth who are black but are adopted by white parents? White youth adopted by black parents?

Selective initiatives essentialize the groups they intend to help. Considering the issue of scale, helping only one group (or a set of groups) can be inefficient, given that people hold multiple identities at any given time.

A computer science department which only provides selective initiatives to women will wind up neglecting other underrepresented groups – who may need the help more than some (but not necessarily all) of the women. It is easy for departments to focus resources on visible minorities such as women and racial groups; invisible minorities tend to lose out on the selective initiative identity politics.

Universal diversity initiatives as a result have the potential to save resources. For example, rather than create a women-in-CS mentorship programme, an aboriginal-in-CS mentorship programme, a deaf-in-CS mentorship programme, and a trans-in-CS mentorship programme, a CS department can implement a mentorship programme for all students. Minority-member students can be stealthily matched with minority-member mentors. Not only are the visible minority groups aided, but the invisible minorities – especially those without faculty advocating on their behalf in department decision-making – are aided as well.

#### 6.1.5 Resources Needed

On the note of resources, fiscal slack can be a necessary (but not sufficient) condition for policy innovation [40]. While universal initiatives are typically cheaper than selective initiatives, they often require a greater upfront cost: this can be politically difficult.

Indicated initiatives require low levels of 'governmental capacity'. This is defined as the ability of the initiating level to implement a policy, and the target to meet its requirements [40]. As professors have a great deal of autonomy over their teaching they have high governmental capacity in their own classrooms. On the other side, universal and selective initiatives suffer from the need for greater governmental capacity.

### 6.2 Leverage and Scale

Unlike targetedness, the amount of leverage an initiative has does not have a direct relationship with scalability. Leverage doesn't directly relate to who or how many people are affected/involved.

Leverage is about having *lasting* change: higher leverage changes are more likely to be sustained over time. The relevant scale here is not scaling over a population, but scaling over time.

Systems are notoriously difficult to change. Policymakers have long noted that enacting a policy doesn't mean it will be implemented as desired ('fidelity') or be sustained as future policies are brought forward [52, 18]. Higher leverage changes are more difficult to enact but they are more likely to stick once made.

Many selective and indicated initiatives are based on the assumption that a group needs special help. This is part of a paradigm that the *problem is the group itself, rather than what the greater system is doing to that group* [8]. This reflects a paradigm well-documented in the women-in-science literature: whether to change the women or to change the system [30].

Low-leverage changes are easy to understand and explain. They're easier from an optics point of view to work on; they provide a concrete change that one can focus on or take credit for. High-leverage changes are harder to pinpoint, especially when the changes are happening. The path to changing a curriculum is more evident than the path to changing a paradigm, and as a result is easier to rally resources around.

Low-leverage changes are also more easily co-opted by agents with other agendas. Sociologists have repeatedly documented cases of "false change": low-leverage change with little effect used to give a false sense that progress is being made, to stave off more radical change [2].

#### 6.2.1 The Goldilocks Zone

Trying to change just a paradigm of a system is a difficult, if not impossible, task. Instead, queer theorists have referred to the need to start by making change by a *Goldilocks process* [16]. This involves starting with medium-leverage changes, then eventually switching to high-leverage changes.

This process is reminiscent of Vygotsky's notion of *Zone of Proximal Development* [50] but on a system-wide scale. You challenge the system with changes which are at the

periphery of what is possible[2], and once you have your 'foot in the door' you continue shifting the system towards your goal.

Psychologists have documented that behaviour affects attitudes [43]. If a department starts making some medium-leverage changes, then people may engage more with the need for these changes, and become more inclined towards high-leverage changes.

#### 6.2.2 Resources Needed

Making high- and medium- leverage changes to a system requires an understanding of the system. Information can be one of the most vital resources needed for high/medium leverage changes.

Counterintuitively, high-leverage changes can be the cheapest to implement: it costs very little *money* to change a paradigm, but to change your TA-to-student ratio is likely quite expensive. But while high-leverage changes may need fewer fiscal resources, they need much more political capital and governmental capacity. Changing the goals of an organization requires a great deal of political support – and lack of organized opposition.

Institutional context hence becomes vital here: how is formal/informal authority allocated amongst policy actors, and how are decisions made [40]? For example, one CS department could make decisions through committees; another through consensus-building. Consensus-focused departments are vulnerable to having policies blocked by professors who oppose the changes. This makes it harder to pass policies that have opponents – meaning that high-leverage changes are harder to make. But a consensus-focused department where everybody is brought on-board for a high-leverage change means the change is more likely to last.

### 6.3 Policy Space

The concept of *policy space* refers to the cumulative effects of previous policies, and how they shape the creation and implementation of new policies [4, 40]. Policies do not live independently, and cannot be thought of independently. Policymakers considering new diversity initiatives in their jurisdictions need to consider how new initiatives would fit into existing systems and interact with existing policies.

If a department or a professor has been used to making changes with a certain level of leverage/targetedness, they are more likely to stick to that level of leverage/targetedness [40]; alternative approaches may be too unfamiliar to them.

Research on how education policies are scaled up has documented multiple stages to the process: first demonstrating a proof of concept that the initiative can be implemented at all, then evaluating that it works, then showing it works in several other contexts, *then* scaling it up and refining it [39]. Frequently, context is the most important factor: many well-intentioned policies have 'failed' for neglecting the impact of context [52].

## 7. DISCUSSION

A couple papers ([15, 12]) have been written asking the question: why are there more women in other STEM fields than CS? Cohoon attributed the difference between biology and CS to the following: biology faculty have more favourable attitudes towards female students, spend more time mentoring students, and feel more of a shared responsibility for student success [15].

Selective initiatives in biology are relatively rare compared to computer science. Instead the biology faculty have goals more focused towards undergraduate teaching, and the greater buy-in toward teaching has resulted in faculty doing more indicated work.

Other work looking at the differences between STEM fields also points to the paradigms in CS teaching as problematic – the collective belief held by CS educators that ability to perform in CS is fixed [27] is attributed to lower diversity [35].

### 7.1 Diversity

Our paper has focused on gender diversity in CS. It must be noted that many other facets of diversity exist: race, class, disability, sexual orientation, gender expression, etc. Different underrepresented groups in computer science have differing reasons for their underrepresentation – and individuals belonging to multiple minority-groups face interaction-effects from the multiple biases they encounter.

At the same time, many of the reasons non-gender minority groups are underrepresented in CS are the same as women: lack of encouragement [28], lack of prior exposure [37], stereotypes [49], hostile attitudes and biases [42], lack of role models [42], not knowing the 'hidden curriculum' [38], and not being part of the 'old boys' network' [54].

In many ways, the culture in computing discriminates against those who do not fit the stereotype of the white/Asian male 'nerd': even white male 'jocks' have reported feeling out of place in the CS classroom [38].

Gender has received the lion's share of the diversity research in CS education, and as a result we found it most appropriate to focus on it for this paper. Given the large number of women-in-CS initiatives to draw upon in creating our framework. Since we adapted very general frameworks to do so, we believe our framework will appropriately transfer to other diversity initiatives.

### 7.2 Limitations

Our conceptual framework provides some insights to the properties of different diversity initiatives, but the work we used to put it together has not been without critique. Meadows' approach to systems thinking ignores issues of historicity and power. While we included the concepts of political support/opposition and policy space in our discussion of resources and constraints, these still leave historicity and power in the background of our analysis.

Both the USI model and the leverage points fall into the structuralist approach to sociology, which traditionally ignores or backgrounds issues of historicity and power. Given the importance of power in diversity issues, future work in examining CS education policy may find it useful to foreground historicity/power – some alternate approaches could have been to use the concept of co-construction [18].

We chose a structuralist approach in this paper because it gives directions forward. While poststructural approaches are useful for exposing the reproduction of inequalities in organizations, they can give very little in terms of ideas for what to do about them. Structuralism gives concrete ideas for educators: *let's try a higher-leverage change; how about something selective? etc.*

### 7.3 Future work

Policy research, especially at scale, has a difficult time of comparing two policies: you can't necessarily scale up both, and you can't scale them both up on the same population. Experimental research becomes infeasible: policies hence need theoretical backing for scaling up, and research on relevant contexts.

Having developed a conceptual framework for classifying diversity initiatives, and presented some preliminary evidence of the importance of high-leverage changes, our next step is to conduct a mapping study of the CS education literature. It appears that much of the literature focuses on

[2]Political theorists refer to the scale of what is possible as the 'Overton window' [46].

the indicated and selective initiatives; universal initiatives appear underrepresented.

Future work is needed to look at the *micropolitics* of how CS departments make decisions on diversity initiatives. Micropolitics refers to the study of politics in organizations [3], and provides valuable insight for why schools and universities favour particular policies over others.

Existing papers on large-scale efforts, such as at HMC [1] and CMU [38], focus on describing the medium- and low-leverage changes that were enacted – rather than the high-leverage points or the context of their institutions. Furthermore, both works fail to describe the micropolitics of their organizations, only providing short and vague messages like "have a champion" for proposed policy changes.

Within the CS education literature, a paradigm of positivism can easily be spotted in papers on diversity. In Maria Klawe's account of the changes at HMC [34] she boasts that "other institutions can easily replicate" HMC's successes in attracting women into CS. We find this overly optimistic, as it ignores the cultural factors at HMC which made these changes possible – factors which include a president like Maria Klawe!

Klawe gives CMU and UBC as other examples of success, but again ignores cultural factors there. Most CS departments feature professors who care about diversity – but their resources and constraints may not favour the changes that worked at HMC. Context is a vital part of understanding what made a diversity initiative 'work', particularly the relevant existing resources and constraints.

In order for other institutions to understand *how* to make changes like at HMC and CMU, we need research on how policy actors navigate the political waters to enact change. This political knowledge is vital for scaling up.

As CS educators strive to make widespread changes to the demographics of their classrooms, we need to think about how to transfer and scale up the findings from the existing CS education literature – much more can be done to use the tools from education policy analysis in this research area.

### 7.4 Take-homes

The purpose of McDonnell and Elmore's paper was not only to reconceptualize policy, but to give conceptual tools to policymakers. They observed that policymakers are often unaware of the range of policy tools available to them, and stick to instruments that have worked for them in the past. McDonnell and Elmore's paper gives a structured way for policymakers to brainstorm policy approaches that would be in their blind spots [40].

Similar to McDonnell and Elmore, this paper gives CS educators a conceptual framework for thinking through what policy alternatives are available to them. When educators find themselves seeking to improve diversity, they have an activity available to them now: to brainstorm a change for each leverage point, and for each level of targetedness. The activity may uncover ideas that educators would not have otherwise considered.

## 8. CONCLUSIONS

While low-leverage, indicated initiatives may be the easiest for a CS educator to start with if they want to make a difference, these initiatives are likely the least effective – and least likely to scale well. Selective initiatives, while popular, present numerous challenges for scaling up; universal initiatives provide greater promise for effective policy at scale.

High leverage changes are most effective long-term, but are difficult to make on their own; medium leverage changes (system behaviour and function) fall into a 'Goldilocks zone': they provide an effective place to start, to start shifting the system toward high-leverage change. CS educators may want to consider what feedback loops, goals and rules privilege majority-group members in their classrooms and CS programmes – and how their undergraduate programmes can be changed to level the playing field for all students.

## 9. ACKNOWLEDGMENTS

A thank you to Greg Wilson for providing a pointer to the USI model, to Nina Bascia for introducing us to the literature on micropolitics, and to Aditya Bhargava for proofreading.The first author has received funding from the Social Sciences and Humanities Research Council of Canada; the third author from the Natural Sciences and Engineering Research Council of Canada.

## 10. REFERENCES

[1] C. Alvarado, Z. Dodds, and R. Libeskind-Hadas. Increasing women's participation in computing at Harvey Mudd College. *ACM Inroads*, 3(4):55–64, Dec. 2012.

[2] N. Aschoff. *The New Prophets of Capital*. Verso Books, 2015.

[3] S. J. Ball. *The micro-politics of the school: Towards a theory of school organization*. Routledge, 2012.

[4] N. Bascia and B. Faubert. Primary class size reduction: How policy space, physical space, and spatiality shape what happens in real schools. *Leadership and Policy in Schools*, 11(3):344–364, 2012.

[5] E. P. Bettinger and B. T. Long. Do faculty serve as role models? the impact of instructor gender on female students. *American Economic Review*, pages 152–157, 2005.

[6] D. E. Betz and D. Sekaquaptewa. My fair physicist? Feminine math and science role models demotivate young girls. *Social Psychological and Personality Science*, 2012.

[7] S. Beyer. Why are women underrepresented in computer science? gender differences in stereotypes, self-efficacy, values, and interests and predictors of future cs course-taking and grades. *Computer Science Education*, 24(2-3):153–192, 2014.

[8] J. S. Brotman and F. M. Moore. Girls and science: A review of four themes in the science education literature. *Journal of research in science teaching*, 45(9):971–1002, 2008.

[9] R. P. Brown, T. Charnsangavej, K. A. Keough, M. L. Newman, and P. J. Rentfrow. Putting the "affirm" into affirmative action: preferential selection and academic performance. *Journal of personality and social psychology*, 79(5):736, 2000.

[10] A. E. Budden, T. Tregenza, L. W. Aarssen, J. Koricheva, R. Leimu, and C. J. Lortie. Double-blind review favours increased representation of female authors. *Trends in eco. & evo.*, 23(1):4–6, 2008.

[11] J. Butler et al. *Gender trouble*. routledge, 2002.

[12] S. Cheryan. Understanding the paradox in math-related fields: Why do some gender gaps remain while others do not? *Sex Roles*, 66:184–190, 2012. 10.1007/s11199-011-0060-z.

[13] S. Cheryan, V. C. Plaut, P. G. Davies, and C. M. Steele. Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology*, 97(6):1045–1060, 2009.

[14] J. M. Cohoon. Recruiting and retaining women in undergraduate computing majors. *SIGCSE Bull.*, 34(2):48–52, June 2002.

[15] J. M. Cohoon. Women in CS and biology. *SIGCSE Bull.*, 34(1):82–86, Feb. 2002.

[16] D. K. Cortese. *Are We Thinking Straight?: The Politics of Straightness in a Lesbian and Gay Social Movement Organization*. Taylor & Francis, 2006.

[17] A. Dalton, J. Patterson, A. Stover, and H. Rilkoff. Suicide prevention in Toronto, 2014.

[18] A. Datnow and V. Park. Conceptualizing policy implementation: Large-scale reform in an era of complexity. *Handbook of education policy research*, pages 348–361, 2009.

[19] N. Ding and E. Harskamp. How partner gender influences female students' problem solving in physics education. *Journal of Science Education and Technology*, 15(5-6):331–343, 2006.

[20] S. Fincher. What are we doing when we teach computing in schools? *Communications of the ACM*, 58(5):24–26, 2015.

[21] N. C. for Women and I. Technology. By the numbers, 2013.

[22] J. W. Forrester. Urban dynamics. *IMR; Industrial Management Review (pre-1986)*, 11(3):67, 1970.

[23] P. Freire. *Pedagogy of the oppressed*. Bloomsbury Publishing, 2000.

[24] K. Garvin-Doxas and L. J. Barker. Communication in computer science classrooms: Understanding defensive climates as a means of creating supportive behaviors. *J. Educ. Resour. Comput.*, 4(1), Mar. 2004.

[25] J. Gastil. Generic pronouns and sexist language: The oxymoronic character of masculine generics. *Sex roles*, 23(11-12):629–643, 1990.

[26] M. Guzdial. Exploring hypotheses about media computation. In *Proceedings of the Ninth Annual International ACM Conference on International Computing Education Research*, ICER '13, pages 19–26, New York, NY, USA, 2013. ACM.

[27] M. Guzdial. Anyone can learn programming: Teaching > genetics, 2014.

[28] M. Guzdial, B. J. Ericson, T. McKlin, and S. Engelman. A statewide survey on computing education pathways and influences: Factors in broadening participation in computing. In *Proceedings of the Ninth Annual International Conference on International Computing Education Research*, ICER '12, pages 143–150, New York, NY, USA, 2012. ACM.

[29] S. Hall. New ethnicities. *Stuart Hall: Critical dialogues in cultural studies*, pages 441–49, 1996.

[30] S. G. Harding. *The science question in feminism*. Cornell University Press, 1986.

[31] M. E. Heilman, C. J. Block, and P. Stathatos. The affirmative action stigma of incompetence: Effects of performance information ambiguity. *Acad. of Mgmnt. J.*, 40(3):603–625, 1997.

[32] C. E. Hmelo-Silver and M. G. Pfeffer. Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science*, 28(1):127 – 138, 2004.

[33] C. R. Kaiser, B. Major, I. Jurcevic, T. L. Dover, L. M. Brady, and J. R. Shapiro. Presumed fair: Ironic effects of organizational diversity structures. 2012.

[34] M. Klawe. Increasing female participation in computing: The Harvey Mudd College story. *Computer*, 46(3):56–58, 2013.

[35] S.-J. Leslie, A. Cimpian, M. Meyer, and E. Freeland. Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, 347(6219):262–265, 2015.

[36] L. Malmi, J. Sheard, Simon, R. Bednarik, J. Helminen, P. Kinnunen, A. Korhonen, N. Myller, J. Sorva, and A. Taherkhani. Theoretical underpinnings of computing education research: What is the evidence? In *Proceedings of the Tenth Annual Conference on International Computing Education Research*, ICER '14, pages 27–34, New York, NY, USA, 2014. ACM.

[37] J. Margolis. *Stuck in the shallow end: Education, race, and computing*. MIT Press, 2008.

[38] J. Margolis and A. Fisher. *Unlocking the clubhouse: Women in computing*. MIT press, 2003.

[39] S.-K. McDonald. Scale-up as a framework for intervention, program, and policy evaluation research. *Handbook of education policy research*, pages 191–208, 2009.

[40] L. M. McDonnell and R. F. Elmore. Getting the job done: Alternative policy instruments. *Educational evaluation and policy analysis*, 9(2):133–152, 1987.

[41] D. H. Meadows. *Thinking in systems: A primer*. Chelsea Green Publishing, 2008.

[42] K. Modi, J. Schoenberg, and K. Salmond. Generation STEM: What girls say about science, technology, engineering, and math. *A Report from the Girl Scout Research Institute. New York, NY: Girl Scouts of the USA*, 2012.

[43] S. Owens and L. Driffill. How to change attitudes and behaviours in the context of energy. *Energy Policy*, 36(12):4412–4418, 2008.

[44] E. Pronin, C. M. Steele, and L. Ross. Identity bifurcation in response to stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology*, 40(2):152–168, 2004.

[45] Z. Richards and M. Hewstone. Subtyping and subgrouping: Processes for the prevention and promotion of stereotype change. *Personality and Social Psychology Review*, 5(1):52–73, 2001.

[46] N. Russell. An introduction to the overton window of political possibilities. *published online at Mackinac Center web site*, 2006.

[47] B. Simon, B. Hanks, L. Murphy, S. Fitzgerald, R. McCauley, L. Thomas, and C. Zander. Saying isn't necessarily believing: influencing self-theories in computing. In *Proceedings of the Fourth international Workshop on Computing Education Research*, pages 173–184. ACM, 2008.

[48] M. Van den Brink and L. Stobbe. The support paradox: Overcoming dilemmas in gender equality programs. *Scand. J. of Mgmnt.*, 2013.

[49] R. Varma. Women in computing: The role of geek culture. *Science as culture*, 16(4):359–376, 2007.

[50] L. L. S. Vygotsky. *Mind in society: The development of higher psychological processes*. Harvard university press, 1978.

[51] D. Wasserman, T. Durkee, C. Wasserman, and C. Wasserman. Strategies in suicide prevention, 2009.

[52] W. Werner. Curriculum and uncertainty. *Social change and education in Canada*, 2:105–115, 1991.

[53] D. S. Yeager and G. M. Walton. Social-psychological interventions in education. *Rev. of Ed. Research*, 81(2):267–301, 2011.

[54] N. Zarrett, O. Malanchuk, P. E. Davis-Kean, and J. Eccles. Examining the gender gap in it by race: Young adults' decisions to pursue an it career, 2006.

[55] N. R. Zarrett and O. Malanchuk. Who's computing? Gender and race differences in young adults decisions to pursue an information technology career. *New directions for child and adolescent development*, 2005(110):65–84, 2005.