

**“The Unexamined Life is Not Worth Living”**  
**Socrates in Plato’s Apology 38a, cited in UWA’s SI webpage**

**Academics’ Comments on Use of the Socratic Index at UWA**

1. *How is the SI used in each Faculty?*
2. *Where it is used, are academic staff provided with information about the distribution and values of SI scores for their discipline that allow them to present their relative performance fairly?*
3. *Will a set of guidelines for use of the SI be produced?*
4. *Is there a set of protocols to determine when and how the SI formula is changed?*

The questions are not about the use of metrics per se or how high standards for research should be. The questions are motivated by concerns that in some contexts at UWA the SI is being used in ways that (a) substitute “management by spreadsheet” for the exercise of academic judgment and leadership, (b), ironically, leave academics without sufficient information to present their relative performance fairly, and (c) leave academics unable to confidently plan how to commit their time and attention. These issues go to the heart of “academic culture” and how to best achieve UWA’s goals.

Members of the Academic Staff Association (and others who were passed on the message) were invited to comment on the use of the Socratic Index and on the answers to the four questions put to the Dean of each Faculty.

I have numbered the responses and arranged them roughly in order of commonality.

**Response 1**

“I like the Socratic Index - it gives some tangibility to the measure of research performance and provides a basis for discussion”

**Response 2**

“Although healthy debate about relative metrics for performance of academics is worthwhile, I wonder if you could explain to me why uwa has so many academics who score close to zero across ever possible metric conceivable except their own self delusion. Unfortunately your opposition to any form of metric has given us the national reputation of an underperforming backwater!”

**Response 3**

“In response to the question of whether Socrates is used “To determine teaching loads”, [it is] ... noted that the Faculty of Arts (among others) uses Socrates data to determine the “broad parameters for the allocation of teaching workload.”

That is not correct. Within the Arts faculty, an individual’s target teaching load is calculated **directly** from SI according to the following formula (from 2014):

$$17.5 + (S_a - S_i) / 4 = \text{target EFTSL load per semester}$$

where  $S_a$  is Socratic Index average for level of staff member,  $S_i$  the Socratic Index for the individual staff member and there is a minimum of 11.5 and maximum of 21.5 EFTSL per semester.

....

.... For Discipline Chairs, it adds another level of complication to calculating who teaches what, a process already difficult because one never knows what enrolments will be for the year, and the funding coming into the School is always for the current year. This use of Socrates has created a sense that 1) teaching is punishment and 2) we are in a zero sum game insofar as a colleagues’ publishing or grant-getting success raises the average in a way that means we are all punished. Because the average is not statistically relevant, and because SI measures change, it does not create a predictable benchmark for ‘performance’ but a shifting goal post and a sense of unpredictability.

**Response 4**

W/Professor Ray da Silva Rosa, UWA Business School,  
President, UWA Academic Staff Association

I read the attachment provided with some interest, and would just like to flag for you one area of Prof Owen's report where there is, I think, some need for clarification. It is claimed on p2 of her document, in the section on teaching allocations, that in the Faculty of Arts the SI is not directly used to calculate workloads. That may be a misapprehension; for each academic in Arts, the workload is calculated according to a formula in which SI is the main factor.

#### **Response 5**

There is one point in Robyn's response re the use of the index by Arts and Engineering that requires clarification. This is the statement that the index is used "to determine broad parameters for the allocation of teaching workload, but not directly ... to calculate workload".

In my view for Arts this statement is inaccurate. What happens with the Arts workload model is that the difference between the individual index score and the Faculty average index score for that level is used directly to determine an individual's teaching load. This is not mere broad parameter usage but direct discrimination between staff based on the same teaching and research contract depending on their index score. Thus if the index is an inaccurate representation of an individual's research performance (e.g. I published a long chapter in a prestigious DFAT book in 2013 that received one index point, whereas if a section of that chapter were to be published relatively unchanged in a certain journal it would receive 4 points, for considerably less research), then that has a direct impact on the staff member's teaching load.

The late Prof Philippa Maddern pointed out in an Arts Faculty meeting with the VC on 19 June 2012 that the consequence of the use of the index in the Faculty's workload model would be to create a vicious spiral for most staff, in which their additional teaching load would prevent them from improving their research output. This is what has happened for the Arts Faculty in the past two years. Indeed, my understanding is that the Faculty index averages for different levels have all dropped, which shows that the use of the index to improve research output in the Arts Faculty has failed. The common and disappointing view is that extra teaching is now seen a punishment for not producing the type of research output valued by the index. This was a matter that Prof Belinda Probert expressed concern about nationally in her report to government early in 2013 about the implications of teaching only positions in Australian universities.

The most glaring consequence of the use by the Arts Faculty of the index to calculate individual teaching loads, not merely for setting broad parameters, is that the discrepancy in actual face-to-face teaching loads between staff in Political Science employed on the same contractual terms has increased, so it's been usual for some staff to do more than twice the amount of teaching that other staff do. This discrepancy is much greater than the difference that Belinda Probert reported exists at University College Loindon between teaching only and teaching/research staff. In other words, teaching only staff at UCL do significantly less than twice the amount of teaching that teaching/research staff there do, but at UWA in Political Science there is a greater discrepancy in teaching loads between staff all on the same teaching/research contract.

#### **Response 6**

As one of the beneficiaries of the SI system ..., I feel a little conflicted opposing it, but can nevertheless see how it is becoming a divisive tool for splitting the academic staff into teachers and researchers, and directing research in a way that does not reflect disciplinary or indeed social imperatives. In any event, I take issue with the statement that "The Faculties of Arts and Engineering, Computing & Mathematics use Socratic data to determine broad parameters for the allocation of teaching workload, but do not directly use the Socratic Index to calculate workload.<sup>2</sup> This is simply untrue: in our school (Humanities), our workload in EFTSL is directly calculated using the SI. Here is how it's done (from an email of December last year, so the figures are out of date, but the method is the same):

"Find your current Socrates score at <https://www.socrates.uwa.edu.au/Admin/AdminDefault.aspx> and then compare it to the average Socrates index for your grade.

W/Professor Ray da Silva Rosa, UWA Business School,  
President, UWA Academic Staff Association

level A 2.04  
level B 10.3  
level C 17.61  
level D 25.51  
level E 33.03

It's the difference between the two figures (halved) that determines your individual workload. For instance, if a level B academic has a Socrates index of 15, she's 4.7 points better than the 10.3 average for her grade. That 4.7 is halved to 2.35, and 2.35 is then subtracted from 2014's normal workload of 35, giving her a workload target of 32.65.

The Socratic average for your grade may have to be adjusted pro rata to take account of your personal circumstances such as a fractional appointment or parental leave during the past six years (the period over which your SI is calculated). For instance, if the academic above had been half-time for the last six years, the Socratic Average for her B grade would be reduced from 10.3 to 5.15, resulting in a much larger workload reduction.

At the very least, the record should be set straight on this at Academic Board.

### **Response 7**

[The Arts Faculty] said "yes" to all questions - i.e. that Arts uses the SI for all purposes nominated. [The Faculty] used the "inform decisions" column and left the "calculate decisions" one blank.

I have some sympathy for this because "calculate decisions" is an unusual formulation and goes to the heart of what you called "management by spreadsheet". Decisions, in point of fact, cannot be calculated; they have to be made. (Calculations can certainly underpin a decision, they just cannot finally do the work of a decision.) Exactly the same distinction applies to Robyn's formulation: "compute satisfactory performance". Again, satisfaction cannot be computed; it has to be decided.

The verb that should have been used in column 2 is "determine" - which when contrasted with "inform" would indicate that the SI will have a material (yes / no) role in making discriminations.

By that test, the SI is certainly used in the "research" component of the performance review as determinative in the Faculty of Arts. It is the "headline" figure and the de facto judgment. In legal terms, it has the status of a prima facie fact. In teaching, SURF occupies the same position. Everything else in the lengthy performance review document is regarded as hearsay, made unreliable by the simple fact that it is coming from you!

....

(In any event, Education have said they do use SI to "calculate" performance; Equally interesting, Science (Chem/Biochem) have said "absolutely not" which indicates that some leaders regard the prospect as anathema)

With regard to teaching loads - Robyn is absolutely incorrect when she says that SI is not used to "directly .. calculate workload", but only provide "broad parameters". Our workload formula begins with a unit target of 33.5 "students" (EFTSLs) which is then added to or subtracted from according to this formula: SI average for level (i.e. "B", "C" in Faculty) - Your SI. If you are above average your target goes down, if below it goes up. It's hard to imagine anything more direct than this.

It is true that after an initial trial yielded some outlandish results - an envelope was introduced to cap extra teaching to a certain level. We have, of course, run the usual arguments that the distribution of SI's is not normal and that it is possible in these small populations to become very much the victim of another's success, and even for a distribution to occur where one stellar individual is above average and everyone else is below. That, of course, is a good outcome from a management point of view!

### **Response 8**

W/Professor Ray da Silva Rosa, UWA Business School,  
President, UWA Academic Staff Association

You may also be interested in this note from a report of the Committee of the Australian Law Deans:

"The issue of a ranked journal list was the subject of heated debate. Four models were produced (UTAS, USQ, UNSW and Deakin). There is no agreement amongst CALD members that a CALD list should be produced but the matter has been referred back to the Associate Deans Research for further consideration with the input of an external consultant. CALD will consider the matter again next year."

Trying to interpret that, it seems that there is disagreement about whether the CALD should even try to prepare such a list and, if so, what that list should look like.

#### **Response 9**

It seems that there is no systematic use of the SI across the campus. I personally make use of the SI when completing my own PDA as I don't fully support the School RPP system. However, I think the SI could be improved as it often has information data lags and in its most recent version it does not provide the same degree of comparative data that it once did to enable you to see how you are benchmarked.

#### **Response 10**

With my recent experience, the only surprise for me in this is that the Socrates data is used to inform decisions about faculty leadership positions. I cannot see the evidence for this in the decisions that have been made, but I guess that's because I don't look much at Socrates, and only look at my own information on it.

The continuing aspect of Socrates that seems incongruous to me is the ranking of journals. I thought that this system had been abandoned by the ERA process, so why is so much weight still being put on it by the UWA, especially as I was recently not even able to find a list of ranked journals on the ERA website. Am I missing something?

#### **Response 11**

I do have a major concern about the use of Socratic SI in promotion. To put my case simply, my SI (at level C) was getting very close to the median SI of Level D in July this year, and I was expecting one more article to get me over the Level D median. Unfortunately, and dramatically, the method of SI calculation changed in August. As you know, the new method intends to reward co-authoring, a practice which I have not been able to conduct due to the nature of my research (or does this really matter?). Anyway, as the result of the new method of calculation, my comparative SI has dropped drastically, falling noticeably behind the Level D's new median SI. I have noticed I am one of the extremely few whose relative SI has suffered from the new calculation method. As a result, I have forced myself to postpone promotion application.

Would you be able to raise a concern like this since I am not allowed to attend the board meeting?

#### **Response 12**

I read the discussion paper attached to your email and do agree that the Socratic Index could be improved. I have recently used it for my PDA and found it did not fully measure my citations (Google Scholar is much more comprehensive), nor did it accurately measure my HDR supervision. It has a lag effect that is inherent in the systems that feed data into it.

Further, it once provided much better benchmarks of performance such as comparison with other LVLE and UWA/Faculty averages across a wide range of indicators.

The use of SI as a guide to individual academic performance has some merit. Yet to use it for hire, fire and promotion decisions as an authoritative measure is highly problematic.

....

This is a problem across Australia and in many countries. It is damaging collegiality and limiting

academic freedom whilst doing nothing to advance the cause of good research and scientific inquiry. I agree that there is a need for a wider range of metrics to assess individual and ultimately group performance. They need to reflect the full range of academic activity that includes the quality, quantity and impact of research. With "impact" the measures should include impact on different targeted stakeholders. So if the purpose of the work is to impact on other researchers then citation measures are appropriate. Yet if the impact is on say industry or policy makers, other KPI can be used. Better measures for teaching and "service" contributions need to be found as well. There are deep flaws in the SURF reporting system with response rates well below 50% making any data unreliable.

#### **Response 13**

The general feedback I have received to what you circulated is that people don't mind Socrates as a tool but they object to the reification of the indices given (i) the errors in the system and (ii) the "work" that becomes implicitly unvalued through omission. One example I have is textbooks I authored. The textbooks get zero points and the revision for the next editions get zero points, but as a teaching-research academic I see generating leading texts that go to second editions as a key academic contribution. A colleague of mine makes the same point about a psych test s/he has authored. It is a real product of an academic endeavor but accrues zero points. I wouldn't mind that there are omissions as long people didn't use the SI to judge my performance **in its totality** or believe that it is a valid indicator of my performance.

#### **Response 14**

One issue to raise would be why the overall Socratic Index is not open to everyone. There is something very uneasy about it being only available to a select few.

#### **Response 15**

Our school has perhaps the unique position to comment on the (poor) use of citation statistics, having the quantitative mastery and expertise to make a valuable contribution to the discussion

*[RdSR's comment: perhaps reassuringly, several Schools across the University believe they are well qualified to assess the validity of the Socratic Index]*

I have asked my HoS for our school to submit a unified statement on the misuse of journal impact factors in the evaluation of research; however, my concerns have not been heeded. (In Socrates 3, a paper whose journal appears in the top 20% of JIFs in the ISI, gains an extra point, and this has created significant perturbations to each academic's ranking). Apart from these sub-issues, there are of course, those you outline: the narrow scope that Socrates takes into account, its use in establishing workload models, the continual revision of the system, the division of publication points by authors (or lack of).

#### **Response 16**

[My] background is in Industrial / Organisational Psychology, so issues of how to assess performance and the consequences of doing so in various ways is highly relevant to our field and I am really pleased that this discussion is going on!

The concerns I had thought about after reading your email were as follows:

[1.] Any use of an objective indicator of performance will almost certainly lead to people engaging in activities that will increase that indicator, but it pays no heed to how people go about this. That it, it opens the door to 'gaming', and potentially harmful but unmeasured behaviours. The famous "On the folly of rewarding A while hoping for B" paper springs to mind here. We already have a vast array of 'objective' metrics out there to assess ourselves on (H-index, citation count, ResearchGate score, i-10, Impact Factors of our pubs, ABDC/ERA Journal Rankings) – do we really need yet another one?

[2.] I can't figure out exactly how it's calculated. Maybe the info is there somewhere but I couldn't

find it immediately. In any case, from your attachment, precisely how it is calculated seems to change from time to time as well. This could mean that people who have done all the 'right' things one year suddenly find their SIs dropping in the next year if/when the formula changed.

[3]. There is a rather depressing section on Socrates on Grants, where I get to see all of the grant applications that I missed out on! The trouble I have with this is that I would have thought that the University would place some value on the behaviour of applying for grants but it seems as though it is only the successful ones that contribute to the SI. Sure, successful ones are better than failed ones, and they should be weighted more heavily, but if failed applications attract no value, but do populate a list that may be discussed at performance evaluation time (and not to mention take a great deal of time to write), I worry that this list of failures will come back to haunt me one day. The end result of this is that we are effectively penalised for applying, but rewarded for succeeding – so given current ARC/NHMRC success rates, it might become better not to try!

[4]. I couldn't see anywhere in Socrates that acknowledged the receipt of non-competitive industry funding. E.g. consulting money that leads to research partnerships or research papers. A lot of the work I do could potentially be funded like this, but unless this type of funding is acknowledged, there is no incentive to pursue this type of funding. Nonetheless, this type of funding often carries with it some extra work. That is, you don't just get to do your research, but have to do stakeholder workshops/presentations/reports etc.

[5]. It's nice to see the little gold stars next to our top 20% JIF publications, but the problem with this list is that the JIF is quite a fickle index, and a publication in the 20% one year might not be in the top 20% next year. Thus an article which gives me 5 SI points in 2014 might give me 4 points next year if the JIF of that journal

[6]. The notion of having to reach the MEAN of the next level in order to be promoted is completely crazy! The logic here is that only once you are performing at a level equal to the middle/average performance of someone in the level above you, can you apply to be promoted to the bottom point on the pay scale of that level. This makes no sense. I would have thought promotions should not be based so much on whether we are *already completing* work at a higher grade than we are on, but instead based on our potential to complete work at the higher level.

[7]. There is an emphasis on things that a person has completed but not much of an emphasis on how well things were completed. E.g. I can see all the units I've coordinated in there, but I cannot see anywhere any record of what % of that unit I actually taught, or what my surf/spot ratings are. Perhaps this will appear later (I am only in my first year of a tenurable appointment).

[8]. There is no consideration for the 'proactive' or 'citizenship' type of work activities. E.g. I really want to spend some time reviewing one of our Masters courses, to make it run more efficiently, but if I do that, there is nothing in it for me in Socrates.

[9]. Socrates seems to place an emphasis on your performance relative to others. The problem with relative performance metrics is that they don't consider how well the whole unit is doing in absolute terms. Thus if a whole unit is being more productive over time, there is no room for acknowledging this

All up, I think most of my concerns really just hinge on my first point!

### Response 17

I am attaching the 'help' information for the new version of Socrates, Socrates III. Those of you who have had a look this week will see that your SI has gone up considerably with this new version (but so has everyone's, so don't get too excited).

#### **Below are some points to note ....**

'In a nutshell, collaboration has been removed as a divisor - that is the principle reason people will have increased. In addition, publications in certain journals have got bonus marks (top 20% of Thomson Reuters categories).

From an Arts perspective, you can no longer really compare the performance of Arts against the University mean/median because the differences between disciplines are now much greater.'

### Response 18

W/Professor Ray da Silva Rosa, UWA Business School,  
President, UWA Academic Staff Association

I have now worked out exactly what they have done that has totally rearranged the rankings on Socrates.

To give one example - a scientist in the University published 32 (!!!) papers last year.

All were in Thomson-Reuters Index lists.

Thus each paper was allocated 2 Socratic Index points.

Unlike in previous years, however, these are not now divided pro-rata; rather, in the particular case I give, I noted some papers each with 14 authors, yet the individual was allocated the full 2 points as if he had written the paper on his own.

The other new thing is that they now give an additional point to anyone with a paper in a Thomson-Reuters list that is ranked in the top 20 per cent for impact factor in the 'area'.

That is nuts in many inter-disciplinary areas.

In Education, for example, it means that publishing in many of the Science Education journals get the bonus point as they have a 'high' impact, but you are stuffed if you are in philosophy or history of education (and other areas) as they do not have high impact factors.

I have listened to the arguments already that the aim of Socrates is to change behaviours - I am just wondering how to place philosophy and history of education papers in Science Education journals, or perhaps nursing papers in physics journals.

### **Response 19**

Just to let you know the new Socrates has gone live.

It is radically changed and social sciences and humanities have been gutted - in the case of Education, staff have all been pushed down the list about 100 (really) places.

Clearly there are now new criteria yet again, but unlike on the old website, these are not made explicit any more - so bad enough to be pushed way way (way) down the list, but not to have the criteria public on the website is the last straw.

Demoralising is hardly the word for it.

### **Response 20 - a series of emails**

To Office of Research Enterprise, UWA

Hello,

I'd like to make sure that the \$10,000 grant I won from the [Industry Centre] is listed here. I was the lead investigator, and the grant was awarded for 2014.

Response from Office of Research Enterprise 14 July 2014

*I am sorry, but only centrally administered grants recorded in InfoEd (the UWA Research grants Management System) are imported into Socrates and count towards the Socratic Index (SI).*

*We are planing on introducing some functionality later in the year that will allow you to record your other grants in your profile, but these are unaudited and will not count towards the SI.*

Thanks for letting me know. Do you know if is there any rationale behind this policy, or is it a case of policy by accident?

As it stands the policy would seem to disadvantage staff who find smaller grants, and benefit those who receive larger ones -- surely the people who need less support. It would therefore seem to run counter the the university's aims and be due for correction.

Response from Office of Research Enterprise August 1, 2014

*Yes, there's a policy behind this, which is in line with creating drivers to improve the university's overall research performance.*

*Socrates is a measurement of research performance: ergo the more you publish or earn (in terms of grants) the higher your score. We've tried to include additional quality drivers as well (such as winning Australian nationally competitive grants as recognised by the Dept of Education in HERDC).*

*The use of Socrates for workload modelling occurs at the local level, and obviously a large part of promotion is based on research performance.*

#### **Response 21**

I have located papers by Anthony Van Raan that deal with the problematic nature of many of the measurement systems that are now driving our daily professional lives. The attached paper is one that despite its publication in 2005 remains highly relevant today.

The following extracts from the paper's conclusions are noteworthy, I have highlighted some particular sections for emphasis:

*"From the above considerations we conclude that the Shanghai ranking should not be used for evaluation purposes, even not for benchmarking. Several aspects of the Shanghai methodology have to be analyzed in more detail. For instance, the number of Nobel laureates as a partial indicator and factor in the ranking procedure. Also here .affiliation. is a serious problem. A scientist may have an (emeritus) position at University A at the time of the award (which seems to be the criterion in the Shanghai study), but the prize-winning work was done at University B. The 1999 physics Nobel Laureate Veltman is a striking example (A= University of Michigan, Ann Arbor; B = University of Utrecht)."*

*"Rankings such as the Shanghai one are part of a larger problem in the science evaluation circus. Quite often I am confronted with the situation that responsible science administrators in national governments and in institutions request the application of bibliometric indicators that are not advanced enough. They are aware of this insufficient quality level, but they want to have it .fast., in .main lines., and not .too expensive.. Many of the problems mentioned by WEINGART (2004) such as the .easy. use of journal impact factors for research performance assessments of individuals scientists or groups simply should not exist: there are already better indicators for quite a long time. But still heads of institutions demand their librarians to do .quickies., i.e., rapid and, particularly, cheap .evaluations. with help of standard journal impact factors. The fault of these leading scientists and administrators is asking too much and offering too little. The responsible persons do not want to pay a reasonable amount for a study of better quality. They do pay a considerable amount for the data, but want to have a competent evaluation study based on these data for a small fee: .the data are already there, so please press the button. Thus, it is not so much the commercialization of the monopolist data producer ISI that makes the problems. These heads of institutions, government administrators and policy makers are the first to blame that the intermediary research groups that hitherto cleaned the crude ISI data, prepared the data for the construction of reliable bibliometric indicators and developed the competence and skill to interpret the indicators (WEINGART, 2004) are being squeezed out of the market."*

#### **Response 22**

*One member advised they wished to be identified in their response*

17 Nov 2014

I feel that the Socratic Index has been used to create a very unfair working environment in the Arts Faculty. As someone who will retire at the end of this year at the age of 70, I have been aware of the limited time and resources I have had to fulfil my career goals and ambitions. I do not think that the Socratic Index reflects my contribution to Anthropology, nor has it reflected the work I have done as an anthropologist and at UWA. At the end of the year I will be leaving my colleagues with regret; I will be leaving UWA with resentment.

W/Professor Ray da Silva Rosa, UWA Business School,  
President, UWA Academic Staff Association

Best wishes, Victoria Burbank