

REPORT OF PROCEEDINGS BEFORE

**COMMITTEE ON THE OFFICE OF THE OMBUDSMAN,
THE POLICE INTEGRITY COMMISSION AND THE CRIME
COMMISSION**

**INQUIRY INTO PERFORMANCE MEASURES AND
ACCOUNTABILITY OF OVERSIGHT AGENCIES**

At Sydney on Monday 20 May 2013

The Committee met at 11.00 a.m.

PRESENT

The Hon. C. Cusack (Chair)

Legislative Council
The Hon. S. Mitchell

Legislative Assembly
Mr L. J. Evans
The Hon. P. G. Lynch
Mr R. J. Park

DAVID SMITH, Professor of Management Accounting, Monash University,

TIMOTHY PRENZLER, Chief Investigator, Australian Research Council Centre of Excellence in Policing and Security, Griffith University,

KERRY HUMPHREYS, School of Accounting, Australian School of Business, University of New South Wales, and

SURESH CUGANESAN, Professor of Organisational Control and Performance, University of Sydney Business School, sworn and examined;

MANDY CHENG, Associate Professor, Associate Head of School, Australian School of Business, University of New South Wales,

ALDONIO FERREIRA, Senior Lecturer, Department of Accounting and Finance, Monash University,

ROBERT WALDERSEE, Executive Director, Corruption Prevention Division, Independent Commission Against Corruption, and

ALEXANDER BROWN, Professor of Public Law, Griffith University, affirmed and examined:

CHAIR: I declare the hearing open. Welcome everyone. My name is Catherine Cusack and I am the Chair of this Committee. The Committee on the Ombudsman, the Police Integrity Commission and the Crime Commission has the function of monitoring and reviewing the Ombudsman, the Police Integrity Commission and its Inspector, and the Crime Commission, its Commissioner and Management Committee and Information and Privacy Commissioners. In keeping with this function the Committee resolved to inquire into and report on ways in which performance is measured and reported on by those agencies. Having reviewed the submissions and held a public hearing on 11 April 2013, the Committee believed it was important to speak to academics who had experience in considering issues of performance measurement in the context of the public sector. The Committee is particularly grateful to you all for making time to share your expertise and knowledge and thanks you all for appearing before it today. I am advised that you have been issued with a copy of the Committee's terms of reference and also a brochure entitled "Information for Witnesses Appearing Before Parliamentary Committees". Have you all received the information for witnesses appearing before parliamentary committees?

Professor CHENG: Yes.

Dr WALDERSEE: Yes.

Professor SMITH: Yes.

Professor PRENZLER: Yes.

Professor CUGANESAN: Yes.

Ms HUMPHREYS: Yes.

Dr FERREIRA: Yes.

Professor BROWN: Yes.

CHAIR: I ask each of you for the *Hansard* record to state the capacity in which are you appearing today.

Professor CUGANESAN: I am appearing as Professor of Organisational Control and Performance with experience in performance measurement in the public sector.

Ms HUMPHREYS: I am an accounting academic in the Business School at the University of New South Wales. I am appearing with research interest in managers' use of the balanced score-card and other forms of strategic performance measurement systems.

Dr FERREIRA: I appear as Monash University staff. I have interest in the area of performance management. I developed a performance measurement framework on issues and properties of performance measurements.

Professor BROWN: I am Professor of Public Policy and Law in the Centre for Governance and Public Policy at Griffith University in Queensland, and also I am a member of the Board of Directors of Transparency International Australia. I am probably appearing primarily because, in addition to those roles, I was lead author on a national integrity system assessment of Australia that was conducted and completed in 2005 in which we looked, at least a little, at the need for a sensible framework for performance measurement of integrity agencies. I guess in that capacity I have given evidence to some committees in previous years.

Professor CHENG: I am an Associate Professor in the School of Accounting, Australian School of Business at the University of New South Wales. I have some research expertise in the area of developing performance measurements systems and how these systems may influence decision-making and behaviour. Even though I am a staff member at UNSW, I am probably not representing the university. I guess I am appearing more as a private citizen with some expertise in the area.

Dr WALDERSEE: I am the Executive Director of Corruption Prevention at the Independent Commission Against Corruption. I am appearing here in that role but also as an ex academic in management.

Professor SMITH: I am Professor of Management Accounting in the Department of Accounting and Finance at Monash University in Melbourne. I come here with research interests in performance measurement and also I guess a little practical experience in the relevant area. In 2000 I was a member of the Audit Review of Government Contracts Secretariat in the Department of Premier and Cabinet in Victoria, where we were responsible for reviewing a number of the contracts entered into by the Kennett administration. I oversaw evaluation of the private prisons contracts performance measures.

Professor PRENZLER: I teach in criminology at Griffith University in Brisbane. My main work is in the Centre of Excellence in Policing and Security, where I manage a program of research called Integrity Systems, which is concerned with identifying best means to understand and prevent police misconduct. We also take a broad view across the policing industry. So we also look at private security. I have done work previously on this topic in association with A. J. Brown's work and publications.

CHAIR: I especially welcome visitors in the gallery from Bougainville and Solomon Islands parliaments, who are here to observe our Committee hearing today. I understand you all received the wonderful summary paper the staff produced regarding evidence the Committee has received so far and the issues we are examining. Professor Prenzler, in one of your articles you talk about police oversight agencies as, in effect, citizens' watchdog bodies. Is that the purpose of these watchdogs?

Professor PRENZLER: It is a debatable point. There are people who work in these agencies who disagree with that. For example, they would deny that citizen complainants are in any way their clients. So there is that view where they take a very detached view. My interpretation of looking at these agencies around the world is that their main work comes from citizen complaints, sometimes from police complaints as well. In that regard they are very much a democratic organisation. They would not be there if there were not complaints from members of the public against police officers. As well as that, we know that complaints against police really are just the tip of the iceberg of dissatisfaction with police. In many cases, for every person who makes a complaint there are at least another 10 people who have had an adverse experience with police and who think, "I would like to make a complaint" but they do not for various reasons. In that regard, complaints often are very representative of real problems between citizens and police in the community. In that sense I think they are very much, or they should be, representative of citizens.

CHAIR: It is good for government to have someone who can receive those complaints.

Professor PRENZLER: That is for sure.

CHAIR: I will go round the table and ask each of you what you believe is the purpose of these watchdogs.

Professor CUGANESAN: My experience probably has been more with law enforcement agencies rather than oversight agencies. My perspective is really about the collective view, and the collective view is that a law enforcement portfolio should be responsible for being efficient, for being effective, but in terms of how there is certainly a need for transparency and accountability. At the portfolio level, if we take those issues around the outcomes and the effectiveness and efficiency, and the how to in terms of transparency and accountability, the oversight agencies have a more critical role in transparency and accountability to the Government.

CHAIR: Do you see them having a role in driving cultural change?

Professor CUGANESAN: In partnership with the agencies they oversight, definitely.

Ms HUMPHREYS: I agree about the importance of efficiency and effectiveness in meeting the needs of their constituencies, making sure they are taken into consideration. I guess that is where the performance measures come into play, establishing where greater accountability is required to achieve those two outcomes.

CHAIR: Can I ask who their constituencies are?

Ms HUMPHREYS: That is a key issue. In establishing performance measures, as this Committee is looking, I think that is a key point to establish exactly who those groups are, and I think it differs for each of the organisations to whom they are accountable.

Dr FERREIRA: The main role of the Committee, of course, is one of accountability but also it is one of ensuring that the use of public resources is being well put to the service of the community that it is trying to serve. These agencies with their oversight role, in a way, they are ensuring that the police and other entities are out there dealing with the community. There are proper ways of making use of the people and the infrastructure available to them. To me it is a twofold issue: on the one hand demonstrating to the community what is being done, but also ensuring the resources that are available are being used in a way that is satisfactory to the wider community.

Professor BROWN: It is a very good question. The purpose of the agency should be defining the framework or the way in which their performance is measured. It is not just an esoteric question; the purposes should flow directly through to that framework. To my mind—this is not inconsistent with anything that has been said—the purpose of the oversight agencies is to support and ensure the integrity of the institutions they are overseeing, and that is a subset of ensuring public confidence in the integrity of the Government as a whole because of the nature of the institutions we are talking about.

The challenge is that integrity breaks down into a number of different dimensions. I think all of those dimensions have to be recognised. There has been a recent review of the Crime and Misconduct Act in Queensland by Ian Callinan and Professor Nick Aroney from the University of Queensland, which is quite controversial and has some regrettable elements, but it also has some good elements. One of the things they do is point out that integrity basically breaks down into honesty, openness, fairness and also to some extent diligence, performance of duty. With that goes efficiency. All these things have to be in the framework somewhere.

The important thing is not to assume that some of those dimensions necessarily cover the others, which is the problem. We tend to talk about accountability being the be all and end all but accountability is very much a process-oriented process, whereas so much of the work of these agencies is about substance and values, not about the process. It is about the end result and the end effect. I think it is important to recognise there are all those different dimensions and then make sure they are all accounted for and they are not subsumed one into another in a way that one loses but that this is an holistic picture.

CHAIR: To ask the question a little differently, many of our institutions have been established as a result of scandal, royal commissions, lawyers and judges in the inspector roles, judges heading up the commissions. Do you think this legalistic model, the way in which these watchdogs have come about, was optimal?

Professor CHENG: First of all, I say I am not familiar with the way agencies come about, because it is not my area of expertise. My personal view is if agencies better establish trust and confidence, and not just external but also internal, so it wants to represent an organisation it is held accountable to the extent that people

trust they are doing the right thing and using resources appropriately, but at the same time you also want to have a system in place that enables people within the agency to have trust and confidence in the measures being used to look at them, to examine them, so they feel they can do the job properly and not worry about justifying every single action to potential external constituents. So, when we talk about accountability, I think it is more important to think about process accountability rather than being held accountable for outcomes. It is important for people to focus on what happened but it is also important to have systems in place to encourage people to think about processes and to be accountable about processes both externally and to the agency.

CHAIR: So, in the case of the police, ordinary police officers should have an opportunity to have input?

Professor CHENG: I believe so, because there is a distinction on reporting outcomes to people interested in the agencies, which is the public, and also having information available and certain performance reports that enable people within the agency to do their job. So for the latter there has to be input from internal of the agency while at the same time recognising that information is to be provided externally.

Dr WALDERSEE: My view is that the oversight agencies tended to exist where the operational system cannot correct itself—where the agency cannot manage its own integrity, its own waste, its own mismanagement. That is where an external agency has a key role. It is that reason that you would see crises as spawning these sorts of agencies because that highlights where the operational agencies are unable to maintain their own systems adequately. I am not sure that crisis is a good way to start because you will tend to have a knee-jerk reaction and it will become legalistic, but crisis will spawn them because it does highlight the gap in the self-regulation of agencies.

CHAIR: Thinking fundamentally about why we have these watchdogs, there are days I wonder whether we overreacted and that the system's failure to protect itself does not mean it was forever incapable of protecting itself, but by establishing a watchdog it looks good and it looks like we have addressed the problem and everybody can move on. But do you feel that integrity needs to be an independent separate agency?

Dr WALDERSEE: I think aspects of it. That would go back to understanding why the system in some situations cannot maintain its own integrity. If you can understand that, you can say a watchdog has purpose x and its goals would be y as a result, but if it comes out of crisis as you describe it is more likely to be an immediate response to that immediate crisis rather than for true organisational capability. There are reasons why organisations cannot maintain their integrity or mismanagement.

CHAIR: Can you mention a few of them?

Dr WALDERSEE: Simple self-interest would be one. Why would you report corruption or expose your own corruption if it will make you look bad as a manager for having had that occur on your watch? It is much better perhaps to get someone quietly out the door. I will speak hypothetically about the police, because we do not operate in that area, but a Police Commissioner needs to retain the support of his officers. If he is also seen at the same time as unduly harsh on people who are misbehaving maybe he will lose that support. There may be reasons for these conflicting goals or conflicting persons self-interest that would cause agencies not to take the steps that they should.

CHAIR: Is that more true of the police than of any other agency?

Dr WALDERSEE: I do not think any agency wants to highlight the fact that it has wasted money, that it mismanages, that it has not looked after children properly, and the goal would be that you do not make a big thing of it. There is a conflict in self-interest between maintaining your position as a senior person in the organisation and maintaining your personal career and what may be the best thing for the agency, and that is the nature of any organisation.

CHAIR: I think Gary Sturgess refers to a blame culture.

Dr WALDERSEE: That would not surprise me. Gary would do that.

CHAIR: Professor Smith, is there anything you would like to add to that?

Professor SMITH: Just addressing the question of whether the establishment of these oversight bodies is optimal in their development, I suspect the answer is probably not, having come from that strong legalistic perspective. But given, as we have already described, that public confidence is a key element of the role that these bodies play, certainly having a legalistic kind of approach might increase public confidence in those bodies. In terms of the development of the accountability mechanisms, it would be more helpful for a broader set of individuals to be involved in the development of things such as performance measures and those kinds of things. If you look at the literature on performance measurement systems, you see that they are very much focused on the idea that those who are being evaluated according to the measures should have some role to play in the development of those measures. So I think a top-down kind of measurement approach is not necessarily going to generate the outcomes.

CHAIR: Are there any examples you know of where there has been that collaboration, with the people being measured being involved in it?

Professor SMITH: Not within the public sector but certainly in the private sector the academic literature is full of illustrations of what you might describe as a more participatory approach.

CHAIR: Perhaps you would like to take it on notice, but can you point us towards a good example?

Professor SMITH: Yes. I will take it on notice but I can provide that.

CHAIR: Professor Prenzler, would you like to summarise?

Professor PRENZLER: Yes. I certainly agree with you that the development of these agencies in Australia and in many other places is the product of scandal, inquiry and crises of confidence in the police. That is a regrettable process. If we had a more scientific consultative approach to public policy we would have had these agencies in place much earlier in our history and they probably would have prevented the build-up of problems in police departments that necessitated expensive and very ugly inquiries and reform processes. I do not think you should ever have any regrets or question the validity of the existence of these organisations. I think there is a consensus amongst all informed persons in this field that they are essential and certainly many people in other countries and other jurisdictions in the world would like to have their own police oversight agency and do not for various reasons.

It is very clear in Australia that these agencies have been very successful in preventing regression to bad practices in policing. Anybody who knows anything about policing knows that it is an occupation that is particularly vulnerable to the development of bad practices, patterns of abuse, a diverse set of forms of malpractice which can easily become habitual and intergenerational. You need a very strong, powerful, independent oversight agency, along with other mechanisms, to prevent these practices developing and continuing. I do not think we are completely on top of the problem though; I think there are still many problems in policing that are not being effectively addressed by oversight agencies. I think that is partly because the response has been partial in many cases, and so I do not think these agencies are really doing everything that should be done to manage police integrity and ensure that police conduct is optimal. I think there is agreement that we do need these agencies, but there is an ongoing issue about exactly what they should be doing to maximise police integrity.

CHAIR: On this key point, people have mentioned that there needs to be collaboration with front-line police in the development of those indicators. Is there anyone who disagrees with that proposition? Is there anyone who wants to add to that proposition?

Professor PRENZLER: I can provide some examples of successful cooperation between agencies and police from various jurisdictions around the world, which I can table later.

CHAIR: That would be great.

Professor PRENZLER: I would just add that these agencies should be in a kind of binary role. They do need to be adversarial. They do need to take on the police and really have the guts to challenge police, to prosecute errant police officers and to challenge the organisation about inappropriate procedures. But at the same time, where they can, they should try to work together to improve policing and use various sources of information and research to make sure that policing is as good as it possibly can be. I think that is best achieved

through a cooperative approach, keeping in mind that sometimes it will be adversarial and sometimes it will be cooperative.

CHAIR: I think the Police Integrity Commission [PIC] has said to us that it is very difficult to get customer satisfaction from the police in terms of their popularity as an organisation, which is one of the challenges.

Professor BROWN: In terms of some of those basic questions about is it true that these bodies tend to be created in reaction to crises and is that good or bad, or are they too legalistic and is that good or bad, I think that part of the challenge here is that we are actually talking about quite different types of agencies, not all of which have been created in those sorts of environments and not all of which operate according to a legalistic model. That has direct implications for performance assessment as well, because they are very different beasts. I think in the important work that the Committee is doing it is really important to recognise that there is a need for a more consistent—hopefully consistent nationally—approach to measuring the performance of these types of agencies, but they are actually very different types of agencies.

For example, the Ombudsmen's offices generally were not created in response to particular crises; they were created in response to sort of a policy shift that was recognised that small crises and problems fed into over a long period. Similarly in some jurisdictions even some of the other agencies, like anti-corruption agencies, have been created after a fair amount of evaluation as to what the gaps are in the system or not. I think it needs to be said that New South Wales has got a particular history of crises and establishing bodies in response to crises that is probably more dramatic than any other jurisdiction actually. That is not a criticism of New South Wales, but it is just a reflection—

Mr RYAN PARK: You are from Queensland.

CHAIR: I think it is a criticism.

Professor BROWN: The other point that flows from that is that I think there is a real risk that sometimes the agencies will operate on a legalistic model, and that can trap them. That has been one of the criticisms in Queensland. The Crime and Misconduct Commission is very lawyer bound in the way it goes about its business rather than being investigation driven. But some of the other agencies such as Ombudsmen's offices, although they have to have lawyers and be very legally conscious, they basically operate quite anti-legalistically and inquisitorially and flexibly and focusing on the substance of the fairness in a common-sense fashion rather than a legalistic fashion. I think those differences are actually quite important.

They also are especially important differences for performance measurement, because not all agencies should necessarily be complaint driven or complaint responsive in the same way that other ones are. I think this is one of the biggest problems of our integrity agencies as a whole. This is identifiable in Queensland as well but also in other jurisdictions. Chris Wheeler, the Deputy Ombudsman, whom I am sure you all know very well, has done some excellent work that I can point you to if you do not already have it that was published in the Australian administrative law institute forum last year. It was mapping out the differences in the nature of the work of, say, an anti-corruption body and, say, an Ombudsman; one being a complaint-handling body and the other an anti-corruption body being actually not really established as a complaint-handling body. The risk is that if agencies just become complaint responsive and turn into complaint-handling bodies when that is not really what they were established to do then that opens up other gaps in the system. The way in which their performance is measured will influence—either reinforce or detract from—whether they are actually fulfilling that core business or not.

CHAIR: They are almost operating on two levels: the case-based complaint, which needs to be managed by somebody, and the wider public interest, which is to enhance integrity—whatever that might mean.

Professor BROWN: Really at an operational level I think the distinction is between being a complaint-oriented body, where you have an obligation to respond to complaints, where people have the right to make a complaint and the resolution of those complaints to the satisfaction, if possible, of the complainant is actually part of the legal mandate and requirement of the institution, which is what Ombudsmen's offices are there for really. That is as opposed to an anti-corruption agency or a misconduct agency where complaints may be an important part of the business because they are a source of intelligence and complainants may have some justifiable entitlement to some information or some outcome. But if those agencies just wait for complaints as their sole source of intelligence or information they will turn into complaint-handling agencies like an

Ombudsman rather than being—and this is the operational distinction—focused on identifying risks, problems, proactively going out and looking for where the problems are by using other sources of intelligence and information.

Mr PAUL LYNCH: Just as a matter of clarification, you talk about the legalistic model. That is essentially the complaint-driven model?

Professor BROWN: The complaint-driven processes or complaint-responsive processes can become highly legalistic. I guess the other way in which we use that term in terms of these institutions becoming overly legalistic is where the investigative processes that they use are legalistic as well. For example, if we assume that an anti-corruption commission should be conducting all of its public hearings in a particular way, the level of legal representation that is required, our definitions of what procedural fairness and natural justice is, if they become too onerous and if the thresholds that define when an agency may do different things becomes too imbalanced then it is an invitation to legalism. This is another problem with these sorts of agencies.

If the discretion in an agency is not flexible enough and independent enough, for example, to decide when it is going to conduct public hearings or whether it is going to make information public or whether it is going to do something in a particular way, then it is an invitation to those who do not like the activities of that agency to litigate in order to constrain that agency. That has got less to do with performance measurement than other issues to do with the risks of overregulating, if you like, by legislation by trying to constrain and define the powers and the discretions of what are meant to be independent agencies.

CHAIR: They are not only litigating in relation to complaints; they are also litigating against each other. Two agencies have spent a vast amount of resources. As the Parliamentary Committee that oversights them, there is nothing that says that was good or bad performance other than we intuitively do not think it is a good thing. There is really no accountability for that.

Mr LEE EVANS: How should key performance indicators be tested for appropriateness?

CHAIR: Dr Cheng, can I ask you about key performance indicators and how they can be tested for their appropriateness and effectiveness?

Professor CHENG: I guess the short answer is that it is very difficult. I do not think measures really can be tested for appropriateness, but what you do have is the literature that points to a series of criteria that you should consider to see whether a measure is considered as appropriate or not. A test will be something in terms of looking at this set of criteria such as whether this key performance indicator [KPI] provides additional information, what we call informativeness, so whether they provide incremental information that helps the users of that information. It will look at things like how precise the measures are and whether or not it truly reflects the underlying performance of the agency and individuals. You will also probably look at its linkages to the strategy plans that are in place for the particular agency, so are there any clear links clearly incongruent with the main goals of the agency. There will probably be a series of aspects that you want to look at in so-called testing the appropriateness. To a large extent you really do need the input also from individuals in the organisation, as was discussed.

Professor SMITH: There is a recent paper in the academic literature by Wouters and Roymans that looks at the use of prototypes in performance measurement systems. It is looking at ways in which in a rudimentary sense the performance measurement systems can actually be tested and refined and go through various iterations. I am also happy to provide that if it is helpful.

CHAIR: Dr Cheng, in developing performance indicators should the collaboration go in with a blank sheet of paper and have people come up with ideas, or are there actual indicators that should be put to people? If so, what would they be?

Professor CHENG: I guess the common model is to start with your strategy goals. When you talk about what is the purpose of the agency you probably start with that: what you want to achieve. When you report the performance externally also you tend to focus on the outcome, so what ultimately you want to achieve with the agency. The best practice would say that for the KPI internally you start from there and you work backwards in a way to see what you need to do to achieve these outcomes and try to assign responsibilities to various staff within the agency.

You want to look at a clear set of cause and effect relationships linking from the performance indicators of the individual staff to the ultimate outcomes you want to achieve. You want to map it out very clearly, because there is also research that shows that if there are perceived to be a valid linkages between what you are being measured on as a staff individual and the outcomes that you see as desirable, that will give you greater buy-in and greater likelihood of achieving successes. I guess in terms of what you do, it is probably not a blank sheet. You may have a blank sheet just to get people to understand what their goals are. Then from there you can work out what would be the most appropriate measures that would allow you to achieve those goals, bearing in mind the criteria.

CHAIR: Have you had an opportunity to look at the evidence that this Committee has already received?

Professor CHENG: I looked at it very briefly. I have to say I have not looked at it in great detail. One of the tools that I personally researched is called the balanced scorecard. That has been mentioned in some of the submissions but there is not a lot of information in the submissions to see how they use the balanced scorecard. If it is used mostly for reporting purposes then it is probably not as effective at managing performance as if it is used to help to build indicators internally.

CHAIR: I do not know whether anyone has reviewed the evidence we have received from the agencies on this issue, but I think this Committee is making an effort to collaborate with those agencies, because we have to report their performance to the Parliament. I cannot speak for everybody, but I frankly do not have a clue whether our agencies are performing better this year than they were last year. I do not think any committee ever has known. All we can do is take what they wish to tell us and be a conduit into the Parliament. So we are trying to develop those indicators. In the conversations we are having with those agencies we are getting denials that they are service agencies. The Crime Commission says it is not an integrity agency, for example. We have heard a list of excuses as to why its performance cannot be measured and why arrest rates are not relevant. I guess I put this forward as a practical example of our trying to collaborate on what the performance indicators could be, but we are not making any progress.

Professor CHENG: When I looked at some of the submissions they listed some of the performance measures they use and they are what we call outcome measures. I agree with some of the submissions that it is very hard to use this to judge performance because the outcomes are so uncertain, there is so much variability and so little standardisation. The alternative to reporting outcomes is probably reporting on the processes. Perhaps the outcomes cannot be measured reliably but there are measures that look at compliance with processes and policies that are indicative of the performance. Noting the number of crimes you detect is ridiculous as an actual target, but how you go about doing that reporting can be more meaningful.

Mr LEE EVANS: My biggest concern is that if we put a key performance indicator on corruption and in a 12-month period the agency does not find any corruption they feel they have not done their job correctly whereas the agency may not have any corruption or it is a clean skin. At the end of the day if you have someone overseeing it to find corruption they would be down to scratching more than just the paint; they would be looking behind the scenes to find corruption where it does not exist. As a key performance indicator they would be pushed to find corruption within an agency where there might not be any.

Professor CHENG: I would argue that those are the outcome measures we do not want to see because that is not a good indicator of performance. We can develop indicators of the processes in place to try to detect corruption in the best way they can independent of whether there is corruption to be detected. The focus of the target will be something that explains how they can perform better in terms of day to day tasks to achieve the outcome rather than the outcome itself.

Dr WALDERSEE: I agree that the problem with the ultimate outcome, if you take integrity as an example, is that for integrity in the police there are many contributing factors. That means you cannot hold any single element, whether the Commissioner, the Police Integrity Commission or Ombudsman, accountable because it is partly driven by each of them. To go back to the issue that many of the agencies arose from crises—I do not think the strategy is clear behind integrity and police—what are the main drivers of that or the objectives within the system and which agency does which bit of it? Until you understand that and make it explicit it is not clear what standard you are going to hold an agency to. What would you expect the Police Commissioner's contribution to integrity to be versus the Police Integrity Commission versus the Ombudsman unless there is some explicit understanding of what we think drives integrity in police and the steps along the way to that? It is hard to hold anyone accountable and have them understand what they are contributing towards.

CHAIR: Can I draw that out a bit? You are suggesting rather than agency accountability there needs to be system accountability?

Dr WALDERSEE: Unless you make explicit your belief of how the system works I do not see what you are going to hold people accountable to. You cannot hold them to the final goal, which is police integrity—using that example—because there are too many factors beyond the control of any one agency. If you lay out that integrity is a function of exposure of tone from the top, of strong investigative processes—I am hypothesising—but if they are laid out you can say those things belong logically with the Police Integrity Commission or Ombudsman and the other group belongs logically with the police. That way you have coherent objectives that a single agency can be held accountable for. You cannot jump all the way through to the end point because there are too many factors involved.

I think what often happens is that in the absence of making explicit the belief on how integrity is brought about we end up with a lot of activities, such as if we expose things or educate people that might help but it is in the absence of any really well thought out understanding. That probably sits with this Committee to think about. In policing there is community policing versus the zero tolerance tough law and order approach. Both of them are conceptual ways to reduce crime but they are very different processes. What you are going to hold people accountable for in the police will depend on what you laid out concerning how to achieve the outcome, and then you can set accountability. In the absence of that there is flapping around.

Professor CUGANESAN: Can I add to that point? We can draw lessons from the Commonwealth Government space where there are numerous Auditor-General reports on the problems in measuring outcomes and the challenge is that outcomes are predominantly driven on a portfolio and across portfolio collaboration basis. If we use the problem of drugs as an example, it draws on health agencies, law enforcement agencies and corrections. You have multiple agencies across multiple portfolios contributing to reducing harm from issues around drugs. It is not just across a Federal level; it is both Federal and State. One of the key challenges when you look at public sector performance measurement literature is the architecture you need to get meaningful outcome measures. The principle there is portfolio budgets, or individual accountability. That is vertical in terms of focus on an individual agency but outcome achievement requires horizontal accountability and collaboration. That is an issue around outcomes.

Having said that, I think in conversations with agencies you can meaningfully consider—I think this builds on the point—how they contribute to the achievement of an outcome and start to articulate, rather than a single perfect measure, which is never going to happen, they come up with a mix of measures that provides indications of what the agency is doing and then some proxy indicators as to outputs. You mentioned, for example, that you certainly would not want to have a corruption target, because if it decreases or increases you do not know how to interpret that. If it increases are you doing a good job or a bad job? If it decreases are you not looking hard enough or are you preventing corruption? Rather than looking at one indicator you should look at a portfolio of indicators where you may have to say, "Show me evidence of what you are doing in the preventative space."

For those agencies where there is a mix of complaint handling and their own work program you can start a conversation about reactive complaint handling using time-based efficiency and resolution measures, whereas for the proactive space it is working with the agencies for some indicators of the quality of their work. That is not easy and it is a challenge that needs to be done in partnership with the agencies that you are looking to measure. That goes to the point that Professor Smith and Dr Cheng mentioned, which is the notion that coercively imposed measures become legalistic, and ultimately if we are trying to change practices and behaviours there is a raft of literature on how you achieve engagement and buy-in. Consultation is one mechanism.

CHAIR: The New South Wales Government is seeking to inculcate values and ethics as one of the pillars for the public sector and the way they seem to be measuring it is with employee surveys and perceptions. Do you think that is a valid measure?

Professor CUGANESAN: I think the key question is always: Is there a better measure? The only thing I can point to is that there is a trend in the private sector towards greater recognition of the importance of values. The private sector certainly uses a mechanism of surveys of employees, often done independently through organisations such as Hewitt Associates, where they get data as to what employees think and how that reflects organisational values. The key issue in doing employee surveys is the question of independence, whether

employees are able to accurately discharge their views, what happens to the data and whether they feel that confidentiality of responses is protected. In my view it goes not so much to the principle of seeking responses from employees but more how that is administered.

CHAIR: Some of these answers seem very frank to me in terms of a quarter of the workforce saying that they have been bullied. The findings were quite confronting. Do you believe that the trend in employee perception over time is a valid tool as a performance indicator as to how the culture is changing?

Professor CUGANESAN: It could be. I stress the point that the risk is that we look at one indicator in isolation.

CHAIR: I am trying to think of one at the moment. I am open to all ideas.

Professor CUGANESAN: Looking at the perceptions of employees but tracking that to other data such as complaints or even just the intelligence is important. I think the challenge in performance measurement is not knee-jerking to a number but having a conversation and potentially using it as intelligence and a learning tool.

Dr FERREIRA: One of the issues I reflected a little about before coming to the Committee was the notion that we do need to have some form of balance. We have talked about performance in relation to outcomes and processes but I think that is part of a broader process. We need some performance measures looking at inputs such as the effort that is put into finding corruption, looking into processes and how well we are doing. Are we efficiently resourcing the development of inquiries into corruption and things like that? We need some measure of outputs such as the number of people prosecuted. Something to do with outcomes is to me how the broader issue is perceived by the stakeholder. It is a third level and it is intangible. You will need to use a survey to perceive how people see the outcomes of these agencies.

Having a range of performance indicators that look into inputs, processes, outputs and outcomes will bring some form of balance that shows straight off between different areas. If we spend more resources it is possible we would perhaps expect—if there is more fraud—for more frauds to be detected. You can see the connections if you have the whole scope. If you only focus on the output measure, which is the number of frauds that have been detected, then we say it is going up and it is bad. But it is not bad if it was there but it was not seen. It is because we have improved the commitment on the resource side that it is showing up at the other end. A range of indicators from inputs to outcomes would be desirable.

CHAIR: Would staff surveys be a legitimate part of that set?

Dr FERREIRA: I certainly would think so.

CHAIR: All we have at the moment is the amount of media coverage. If there is lots of it it obviously needs more money.

Mr RYAN PARK: The ICAC will be loaded after this.

Professor PRENZLER: I would like to support what has been said about the need for as many indicators as possible. There is a process in research called triangulation, which is slightly misleading because it implies three sources of information but it is actually two or more and the more the better. The answer to your question Mr Evans is: Is there a convergence between the different sources of information? If you go back to your example and you have an integrity agency that investigates complaints and it appears from one set of performance indicators that these complaints have been investigated thoroughly with high quality investigations and they find no corruption and say, "We cannot find corruption", you can compare that against the other sources such as public perception and experience surveys that can be revealing about people's experience of police.

You can then narrow that down to arrestees as another standard measure by surveying people as they come out of court about their experiences with the police: Were you beaten up? Were you given your rights? Did the police try to bribe you? You have the complainant data as well and you can have follow-up surveys with the complainants. You can also survey key informants in the field such as journalists, both defence and prosecution lawyers and legal organisations who work with police, or against police if you like such as youth legal services—those sorts of groups.

What are you hearing from your clients about how police are treating your clients, that kind of thing? You line up all the survey results and if there is a convergence that says there is very little evidence of corruption you think, "This is fantastic and good news", but you might have one source, for example an employee survey, that indicates something else. I am strongly supportive of employee surveys. I have seen a lot of surveys of police. If they are guaranteed that their responses are anonymous you can be fairly confident they will be forthcoming and honest about their colleagues—not necessarily about themselves but certainly about their colleagues. There are a lot of standard surveys of police in the field about using ethical scenarios and also perception and experience surveys: Have you observed colleagues engaging in these types of corruption? If you find that 40 per cent of officers are saying, "We think that fabrication of evidence and manipulation of witnesses is still going on", and the agency cannot find anything then you have some sort of disjunction that needs to be investigated. If they line up you can be fairly confident that the agency is being effective in not finding anything, and that is a good-news story.

CHAIR: Thank you. That is really helpful.

Professor BROWN: One of the key things about those sources of data, whether it is employee surveys or whether it is client, customer or general public types of surveys, we think that the data can be incredibly useful and should be part of the mix. The challenge for the Committee is whether it wants to use that sort of data from the big pool as a performance indicator for the oversight agencies or whether you are going to look for data that is specific to the performance of the oversight agencies. This really goes back to some of what Suresh and Robert said about some of the outcomes being influenced by so many factors across sectors, portfolio areas, across the whole Government, that you cannot necessarily use. For example, the overall rise and fall of public trust in government cannot be blamed on the integrity of the oversight agencies; it may be influenced by too many other things. But you can use the same tools to tap indirectly into performance indicators of the oversight agencies. Part of the skill is in how you do it.

For example, once upon a time—I do not know if it still does—one of the ways that the Independent Commission Against Corruption—and quite a lot of other agencies—measured its effectiveness was to measure whether there was a high level of public awareness of the agency. In fact, that can be misleading. Whether or not people have heard of the Independent Commission Against Corruption or the Ombudsman may not mean anything. The question that measures overall levels of public confidence is that if people had a complaint or concern they would be able to find somebody to go to who can deal with it and deal with it well. That would be a more accurate indicator over time of whether the system as a whole was working or not, but you could not trace it back to individual oversight agencies or integrity agencies. That would be an example of what Suresh was talking about before. It would still be useful, but it would be something that the system or the integrity branch or the group would have to be held responsible for against being collectively responsible.

There are other things that you can do that are much more specific. In our integrity system assessment report we mentioned that the client satisfaction surveys that integrity agencies conduct can be useful and revealing, provided they are not done for public relations purposes. For example, when I worked for the Commonwealth for the Office of the Ombudsman one of the best objective measures that we could identify over time of whether we were performing well or not was the proportion of complainants whose complaints were found to be out of jurisdiction, were not substantiated or upheld who nevertheless said that they believed that the office did a good job, or behaved fairly, or competently. You could measure that over time. If that proportion of people was staying stable or going up you could say things are working well.

Mr RYAN PARK: As someone else said, it is more about satisfaction with the process that the oversight agency went into, not necessarily the outcome?

Professor BROWN: It would be an indicator of a number of things. Potentially you would not necessarily know everything that it was an indicator of, but you could point to it as being a reliable objective indicator that things were working well. If the proportion of people who you would assume would be likely to be unhappy are in fact saying—

Mr RYAN PARK: The opposite—

Professor BROWN: —at some level they are happy, and you can measure it consistently over time and you can compare that between agencies, if you do it consistently then you will have something useful that the agency can use. We used to say, "Hang on a minute. There is a problem, because this proportion of people

who are unhappy is going up and we should be trying to find a way to keep it stable or make it go down." I agree with what Tim said: that you need lots of indicators. The crucial challenge for the Committee—really what we need is a set of national principles for how this is done. There is a question in there about benchmarking, which is very important because being able to compare and contrast in a reliable way, fundamentally, is one of the most useful things in respect of informativeness about the information. It has to be one of the most useful things to be striving for, but it is just very difficult. To do that, part of the challenge is to say, "Here are all the potential and reasonably efficient and reliable pieces of data that you could collect to measure performance. Now let us pick the ones—it might be three, four or five, each on a number of dimensions—that we can trust or that can evolve into something that everybody will trust as being objective measures of something that then can be used to look for changes over time to compare and contrast." That is really the crucial need from the point of view of evaluating how the systems and institutions work in a larger public policy sense as opposed to a management sense of ensuring that the agency is providing value for many in a nitty-gritty sense.

(Short adjournment)

Mr PAUL LYNCH: Professor Prenzler, early in this session you said that you thought the oversight agencies for police had done not a bad job: corruption was better than it had been but there was still some way to go and some things that had to be done?

Professor PRENZLER: Yes.

Mr PAUL LYNCH: How did you draw that conclusion? What performance measure did you use to get to that conclusion?

Professor PRENZLER: It is partly this question of intelligence as well as complaints that are picked up by various sources and reported publicly. If you look at complaints against police, for example, traditional corruption, narrowly defined bribery as well as legal process corruption—fabrication of evidence, verballing and all that sort of stuff—

Mr PAUL LYNCH: Loading up?

Professor PRENZLER: Yes. That is a very small fraction now of complaints. Also, investigative journalists and whistleblowers are not coming forward with a lot of that material. So a lot of that type of organised corruption that occurred in the past—protection rackets and all that sort of thing—appear to have been minimised.

Mr PAUL LYNCH: The best police force money could buy?

Professor PRENZLER: That is right. A lot of that is very much part of history now. I think it is the vigilance of these agencies and the ongoing work that is the best explanation. So when any kinds of little rackets develop they seem to be picked up and stopped fairly quickly. But then we still have problems. I think the number one problem is excessive force by police, inappropriate force, and then there are a whole bunch of problems around inaction, unresponsiveness, incivility, oppressive behaviour and that sort of thing which make up the bulk of complaints and which are revealed through things such as YouTube, video and mobile phone cameras. It is all out there now in a way that it was not in the past. We still have an ongoing problem with police misconduct that is not adequately addressed. I know it is possible to really reduce these problems in policing and we are not very good at it in Australia.

Mr PAUL LYNCH: I do not disagree with anything you have said. That would be very much my general perception about what has happened.

Professor PRENZLER: Okay, great.

Mr PAUL LYNCH: So we are on the same page. I guess the issue I have about it though in this discussion is that none of that really has to do with performance measures. That simply has to do with empirically looking at what happens—your gut feel, almost—and assessing what you see rather than having particular performance measures.

Professor PRENZLER: Yes.

Mr PAUL LYNCH: Do we actually need performance measures? Can we just do the sorts of things you have just done?

Professor PRENZLER: I think those indicators are performance measures in that they reflect on the work of the agency as well as the work of police professional standards departments and police managers themselves. The ultimate performance measure for these agencies must be the level of police integrity as it appears to be from different sources, keeping in mind the question of how responsible agencies are for what happens in the police. But within the sphere of their influence, I think that is a performance measure.

The Hon. SARAH MITCHELL: Earlier we were talking about how we actually determine effectiveness. Catherine referred earlier to the fact that as a Committee our job is to provide these reports to Parliament annually—we hear from the agencies, we have hearings and we take that information and tick that requirement, but it is difficult for us as a Committee to determine the effectiveness based on what we are told. Hence, we have this inquiry. We have spoken about other jurisdictions and a few of you have referred also to the private sector areas. If we are to look seriously at what we need to do in terms of KPIs or these measures we need to know whether there are other examples we could use rather than necessarily reinventing the wheel and try to deal with this ourselves. Some of you will take on notice examples of other jurisdictions or places we could look to for advice or examples, but would anyone like to comment further on that area?

Professor BROWN: I would make a comment that I think, unfortunately, you are in the invidious position where the thinking is that what the New South Wales oversight committees have been doing on this is probably ahead of other Australian jurisdictions. All jurisdictions grapple with the same problem. Certainly in New South Wales, because there has been a stronger and more stable system of committee oversight of the integrity agencies overall for longer than anywhere else really, it actually means that if anybody is well placed to drive the satisfaction of this need for a more consistent package it is actually New South Wales. The exception is Queensland, because it has had what is now the Parliamentary Crime and Misconduct Committee overseeing the CMCs from the word go, even though there have not necessarily been other committees overseeing the other integrity agencies.

The really interesting thing about the recent review in Queensland by Ian Callinan and Nick Aroney is that they have really said the same thing that this Committee is sort of saying: the tools the Committee has to exercise its oversight function of the integrity agencies are not good enough. I think that is probably an accurate description of the state of affairs. I think we have just been waiting. Obviously, parliamentary committee oversight of the integrity agencies is crucial. Having it in the more robust fashion that we have had since about 1990 in a number of jurisdictions for a number of agencies has been a damn sight better than anything else that anybody could come up with. But I think we are now recognising that for that to work well the parliamentary oversight committees really need to have a good package and the proper resources to be able to really know how they are forming their opinions and not just be forced to just accept whatever information is coming from the agencies.

I think that is quite a significant development that really this is part of. I think the crucial thing—we were just talking about this—is that there really is a need for a national look at this; a set of national principles and a national framework taking into account all the differences and distinctions we have been talking about. It does not really matter who takes the lead on developing that, provided it is done. There is enormous benefit if it is done nationally because then jurisdictions can compare one against another. If you look at the Callinan and Aroney report on the CMC, they make an enormous amount out of complaint statistics and try to compare what the CMC handles and does in Queensland with other jurisdictions. But it is very unreliable because there are such inconsistencies in what is a complaint, how is it measured, what it means and what the numbers mean. Really, there needs to be a framework that can resolve at least some of that so that agencies and governments can look across different jurisdictions and say, "Hang on a minute, something good is happening over here."

CHAIR: I am sorry to interrupt, but we have only half an hour remaining and I am trying to work through everything.

The Hon. SARAH MITCHELL: Certainly. Your point is to look at positive achievements by the agencies as well, not only to find deficiencies. If certain agencies are doing things well, that should be acknowledged also. Benchmarking could be something we could continue with what we are talking about. Ms Humphreys, you have an interest in the importance of benchmarking in measuring performance?

Ms HUMPHREYS: Absolutely. It goes back to what Professor Prenzler was talking about earlier: triangulating performance measures. We talked about employee surveys being an important tool for gathering information that we can use to then measure performance. Benchmarking fits with part of that in making sure we have information over a period of time or for a particular agency or comparing different agencies, which again allows us to interpret that performance data.

So what we really want to know at the end is what the performance information is telling us. We do not want to know whether we have more incidences of corruption in a particular year. That is an outcome measure we are interested in but triangulating measures and using tools like benchmarking provides a better source of reliable information along that path. Again, it links into the process and outcome measures that we were talking about earlier. Having process measures gives us another sort of information before we reach the final outcome, which can be much more difficult to measure. Employee surveys fit in as a tool and they are also very easy to benchmark, that kind of information. So it is a useful way of again assessing the reliability of that.

Dr WALDERSEE: Can I just add something to benchmarking?

CHAIR: Yes, I was just about to ask you.

Dr WALDERSEE: In part because this is driving a lot of the work we are doing as the commission and it is also something I did when I was at the AGSM. If you benchmark at a high level, the CMC against the Independent Commission Against Corruption for instance, it is sort of nonsensical because the CMC is in a different State with a different set of problems and a separate set of jurisdictions, and the benchmarks become somewhat useless. If you bring benchmarks down to key issues within the organisation such as complaint handling or how you install an IT system then you are not constrained to looking at similar agencies. You can look at what the private sector does here, what another the Government department does, and you can actually look at using the benchmarking to really drive improvement. One of the mistakes, I think, in benchmarking is to go to too high a level rather than looking at specific issues.

CHAIR: Do the watchdogs in New South Wales discuss these issues about performance management? Do you have contact with the Police Integrity Commission or other agencies that perform similar oversight roles or integrity roles and talk about these matters?

Dr WALDERSEE: Not formally. I chat with the Audit Office and those on regular occasions.

CHAIR: I was interested because it is a shared problem.

Dr WALDERSEE: Yes.

CHAIR: Professor Smith, do you have any comments you wish to make on that?

Professor SMITH: On benchmarking?

CHAIR: Yes.

Professor SMITH: Only really to agree with what Robert was saying about the importance of comparing like with like and drilling down to the process level so, to use an example, you are comparing complaints handling versus complaints handling, so you are comparing apples with apples. The other thing which is a tempting thing to do with benchmarking more generally, particularly if it is supplied in the private sector, is to think about appropriate targets. But this is a sector in which target levels are particularly problematic. One of the perceived benefits of benchmarking is, to a large extent, either difficult to achieve or almost lost because of the sector in which we are operating, because politically identifying targets for certain metrics is going to be either inappropriate or at the very least problematic.

CHAIR: So what is your advice?

Professor SMITH: I think from a political standpoint it is probably important that benchmarking not necessarily be used for setting targets and comparing performance in terms of targets across certain aspects of performance, because of the sensitivities.

CHAIR: But you can benchmark without targets, can you not?

Professor SMITH: Yes.

CHAIR: You can track perceptions over time?

Professor SMITH: You can.

CHAIR: What is the most important thing you think this Committee needs to know? I will go around the table, because everybody has come hoping to accomplish something in their evidence, and I am concerned we might not have asked the right questions of the right people. I would like to get that feedback from each of you.

Professor SMITH: The key thing I would think is just thinking about the integrity of the data being collected. It is important too. It is not something that has been touched on a great deal today. In terms of tracking through what do we measure, arguments around do we measure outcomes or process, there are two key issues. One is data availability and the other is data integrity. One thing I think is important to consider is: We are getting this data; how reliable or valid is it? If I am thinking of something, for example, like a number of reported incidents through a survey, as we were talking about before, it is important to consider the constraints placed upon those people who answer those surveys, if they know those surveys are being used for evaluation—not only in terms of their own willingness to report but also in terms of pressures placed on them to not report, particularly when superiors might know that this data is being used for evaluation purposes. I would like to see a lot of thought given to these issues of data integrity. It is an important thing that we have not discussed so much today.

CHAIR: So you would support its being externally collected?

Professor SMITH: No, I think there should be a mixture. I have no problem with there being a measure of subjective and more objective measures but it is important to think behind the measures and think about what is contained in the data collected.

CHAIR: Professor Prenzler, I might turn to you.

Professor PRENZLER: You want me to sum up?

CHAIR: And I know you were seeking to make some comment earlier. The floor is yours.

Professor PRENZLER: Thank you. I would like you to indulge me and I would like to draw attention to the study I did on this topic 10 years ago, a somewhat similar process to what you are going through. A colleague, Colleen Lewis, and I examined the annual reports of all the oversight agencies at the time in Australia and looked at their performance measures and broke down what they were doing. We also interviewed all the directors about their opinions on performance measures and we compared the results to what the literature said were the best performance measures for these sorts of agencies. Out of that we developed our own model, which was a four-part model, where we advocated four dimensions of performance. There is a bunch of measures that cut across them and informed whether or not performance has been achieved, whether it is high or low and that type of thing.

Very quickly, the first goal was procedural integrity. So when these agencies investigate complaints and process complaints they need to be free of discrimination, bias or even corruption themselves. There are different ways to evidence that. One way is through case study file audits, for example, where you have experts go through a random sample of case files and examine the quality of the investigation and any indicators of perhaps bias or neglect or those sorts of integrity indicators. The second one was procedural rigour. That is the thoroughness of the investigation: were all the leads and evidence followed up and all witnesses interviewed? How thorough was the investigation? Again, there are measures for that such as case file audits, complainant surveys, surveys of the subjects of complaints, success of prosecutions, feedback from judges or tribunal members—a number of sources of information.

The third area we called timeliness but I would change that now to responsiveness. How well did these agencies respond to the needs and interests of complainants and other stakeholders such as the police officer who is the subject of complaint? One of the biggest problems people who go through this system have is lack of communication and delays. They hate that. This is an area that can also be assessed and performance targets set

using benchmarks and all that sort of thing. The final area we identified was impact on police integrity. Again, that is where you use intelligence, stakeholder surveys and interviews, various sources of information to judge the impact the agency is having on police conduct. With that I would like to add the importance of the idea of complaint reduction as an explicit policy. In theory, when these agencies were first established complaints against police should increase as the public have greater confidence, and the police themselves, because a lot of these agencies are set up to provide a place for police to go to because of the failure of internal systems in the past. Complaints theoretically should go up but as they are addressed adequately and as remedial measures are implemented, as there is a learning process from complaints, they really should go down. There are examples from around the world where complaints against police have substantially reduced and police-community relations have improved. I think ultimately complaint reduction, along with other measures of police integrity, should be goals that are set for these agencies.

CHAIR: In the article you quoted from you referred to case file audits. Can you describe why they are useful as a compliance indicator?

Professor PRENZLER: They are something that oversight agencies usually use to assess the quality of police investigations of complaints, where you have review rather than direct investigation. So you get a panel of experts, say an experienced investigator, perhaps a judge, prosecutor or defence lawyer, a mixture of people who have significant experience with investigations, and they just go through a random sample of files and assess whether this is a quality investigation—if leads have been followed up, forensics followed up and that kind of thing. They may ask follow-up questions. That can be random and it also can be targeted so in higher profile cases they might go through the material. Where oversight agencies conduct investigations themselves, you need to have a system. The CMC claims to have done that at times in the past but they have not reported the results publicly. That is not an expensive or difficult process to set up and that can be a useful learning process as well as a kind of report card on investigative quality.

CHAIR: I thank you for the package of information you provided to the Committee. At the conclusion, if you are happy to table those documents, we will have them incorporated.

Professor PRENZLER: Sure.

CHAIR: Professor Cheng, do you have any comments?

Professor CHENG: Only briefly because of the time, but I think it is important that it is important to remember that—

CHAIR: Please take your time. The idea is not for you to feel under pressure.

Professor CHENG: Right. I guess one thing I would like to point out is that performance measurement is really a performance measurement system, so it is more than just measuring performance. One thing to bear in mind is that as you develop the measures, it has to be communicated that they are not just there to monitor the performance but also to enable you to do your job better. If you want to have many performance measures that is great and it gives you a better picture, but you do not want to communicate it in a sense that we are checking everything that you do. We want to communicate quite clearly that these measures will enable the agency to do the job better. The same with employee surveys: it is not so much checking on whether initiatives are put in place to make employees more ethical but to point out the areas that can be further improved. The system involves feedback and learning as well as just monitoring. If you focus too much on monitoring and checking the people doing a good job it will be very hard for any type of performance measured to work at all. That is the one thing I would like to point out.

CHAIR: If I can add to that, in the Government's workforce survey less than a quarter of employees felt they had feedback about their work either formally or informally. The sense was that people would love to have more feedback about their work.

Professor CHENG: Yes, definitely. If agencies are established in response to a complaint it is very defensive. If the measures are implemented in a defensive gesture really they will not be helpful no matter how good the measures might be themselves. You really have to think about improving the feedback.

CHAIR: And maybe it could be perceived more positively?

Professor CHENG: Yes.

CHAIR: Professor Cuganesan, would you like to add anything?

Professor CUGANESAN: Following on Professor Cheng's point, I have three things written down. The first thing is partnership. The worst thing this Committee could do is to just impose a set of measures on the agencies. It goes to the point that you are really trying to balance an accountability relationship of the agencies with this Committee as well as the usefulness of the performance you are measuring. If we are ultimately looking to change practices the challenge is to get those measures used in the agencies, and going down to almost the level of managers of the frontline, because that is where you see a significant influence in practice change. In the public sector the tendency has been—certainly the Australian National Audit Office reports to this effect—you can have a whole series of measures from an accountability relationship and a significant disconnect between that and an internal management system. This really reinforces Professor Cheng's point about performance from an accountability perspective to this Committee, and other stakeholders, but also performance measurement in order to drive change, better positions and better practices. The first thing was partnership. The second thing that I had here was what I spoke about in response to your question a bit earlier about a start point. You can certainly look to the roles of these agencies as enshrined in legislation, but also their strategic plans, their vision and how they are setting out to achieve their vision. I think that is a useful start point in order to have that conversation as to how they should measure their performance.

Then the third things are really dimension and level. I think you do need a set of dimensions, and Professor Prenzler articulated a number of dimensions: What does it make sense to evaluate you on in terms of effectiveness, and what does effectiveness mean? What are you really trying to set out to achieve? And what is standardised and repeatable so that you can actually have conversations about efficiencies? That might be particular types of complaints handling, which then also speaks to the point about benchmarking at that level, and then whether it is equity or access or integrity or some other sort of dimension of performance. I think all of that needs to happen to facilitate that conversation. I said dimension and level because I think the level of performance should be at some sort of program of activity, because the nature of what is done will define the performance that you are looking for. Then I go back to the point made about maybe there is something around complaints handling and it is reactive and you are trying to be responsive and seek effective resolution within a set number of days.

Then there are the more proactive intelligence-based programs at work, and they require quite different mindsets. Indeed, policing agencies themselves have been struggling with this in terms of trying to evaluate their mix of workflow between the reactive investigations of a reported crime and the proactive investigations of serious and organised crime. It is a challenge for them and I think it is a challenge for this Committee just to think through what is being done, what is the performance that you are seeking to achieve and then what indicators you use.

CHAIR: We asked the witnesses from the Crime Commission if they would set themselves a goal of eliminating organised crime. Their response was that it was hopeless and they do not have that as their goal. They said that it was not realistic.

Professor CUGANESAN: It goes back to that issue about an outcome really requiring multiple jurisdictions. Fighting serious and organised crime requires collaboration between a number of Federal agencies, for example, customs and intelligence agencies. You might look to the Australian Federal Police and the Australian Crime Commission and then of course the local State-based jurisdictions. But then it is also around the courts as well in terms of what they are doing in terms of sentencing and being able to prosecute some of these criminals. I certainly agree that an outcome about eliminating serious and organised crime and holding an individual agency to account for that is incredibly problematic; however, having a conversation as to how you might influence progress towards that goal is a useful conversation, then saying, "Okay, let me actually think about some indicators that are not going to be perfect but will at least provide some insights into your role in this."

Dr FERREIRA: A lot of the points that I had have already been mentioned, but I will nonetheless put them to the Committee. I think to me the key issue to take away would be one to do with the issue of balance and looking at these agencies from a more holistic perspective. The focus on outcomes are of course created because we want to see the bottom line. The bottom line might be some measure or some key performance indicators that have high visibility, but to get at the bottom line there are a lot of things that need to happen

before then. The point I would like to emphasise is the idea of having some balanced view of the organisation and the different elements of this chain of events that happen within that so that the bottom line does occur.

The second point that I have is to do with use and purpose. To me, how the information and the performance measures are to be used is really fundamental. If their purpose is to be used for budgetary allocations—as we have spoken of over the break—that is going to put a lot of pressure on them. It is going to drive a more adversarial perspective from the agencies versus the oversight committee. If the performance measures have more of a signalling role in the sense of indicating areas of concern that the committee is interested in evaluating but not necessarily having a dollar figure attached to it then it is more likely that the agencies that are being oversighted are more willing to embrace the process of let us look at these performance indicators and how can we use them, as Dr Cheng noted earlier, to enhance what we do. To me, a clear definition of the purposes or the intentions for which the KPIs are to be used is fundamental.

Clearly, there may be issues to do with resource allocation. If that can be pushed away from the table you are more likely to get more collaboration from the agencies, because that is going to be their concern. They are concerned that their budget will go away and they will be struggling. To me, it is quite important to have clarity of purpose. What are the KPIs going to be used for? I think quite an important process of a performance management system is to signal areas of concern for the community and for the individuals. To me, that is very important.

CHAIR: Is that something the Committee should be doing?

Dr FERREIRA: I think so. That is the way I would see it. That goes back to the point that Professor Cuganesan was making earlier in terms of understanding what is the broader vision. What is it that we are trying to achieve? What are the big areas of concern? What are our goals and what strategy can we put in place? The performance measures, or the KPIs, are essentially a means to an end. They are ways of determining whether progress is being made in the right direction. It may be used just as a means of assessing progress, and from that perspective it is as insightful for those within the agency as it is for the Committee. If the Committee says that we are interested in corruption or we are interested in efficiency that signals something that is important. There is not necessarily a need to set a target.

For instance, there is this movement called "beyond budgeting" where they specifically do not set targets for specific agencies; they set targets relative to other agencies. Frequently what they do is they create what they call a league table where you do not necessarily say that you have to achieve a certain amount of dollar sales or revenue for your agency; it is how you perform in relation to others. Essentially they produce, on a regular basis, a list of, for example, the different bank branches and how they perform in relation to others. Nothing is said. Everyone understands that if they are on the top of the list they are better; if they are at the bottom they are not as good. You need not say to every agency that your number of reported crimes is not in accordance with other agencies. How they rank on that list in itself has a value of signalling that it is important, but at the same time it lets them understand what they need to do to achieve a better performance. That may be a way forward, which leads to benchmarking and involves looking at other agencies perhaps.

Mr LEE EVANS: It would probably assist with getting agencies to talk to one another and set common goals. A league table might achieve the ultimate goal of having all agencies on the same page and trying to keep their benchmarks high for their service to the community.

Dr FERREIRA: Certainly. Someone was talking earlier about this more systemic view of all these agencies. They are not isolated bodies; they are part of a broader system and they are all operating in a similar space. Why are they not talking to each other?

Mr LEE EVANS: Unfortunately, they are all individuals—out of something from Monty Python.

CHAIR: My final question arises from a submission that was made to us by the Public Service Commission. We received a short submission from them that referred to ongoing reporting burdens on agencies. In terms of us thinking about performance management, there seems to be a default perception that it will simply involve yet more red tape. They have to report on the percentage of their employees who have a disability, are female, or are Aboriginal. There are enormous reporting requirements on an array of agencies. In addition, being public sector organisations, the risk management bars are higher, due process expectations are higher and there are more places to complain about due process. To be perfectly honest with you, I find wading through pages and pages of activity reporting not enlightening. I would be happy to see all of that go. I

am interested in your thoughts as to whether we can have a framework for performance management that helps and is less onerous. I can tell you now that every time we ask this question of people we can see their faces cloud over and they think, "Oh no, not another report."

Ms HUMPHREYS: I think there are two things there. The first is making sure that the measures genuinely link to the objectives and the strategic goals of the agencies. That is what stops you from having to read through pages and pages. In the corporate sector it is sometimes hundreds of pages of performance indicators that do not tie together and do not tell a story. If they link back to the strategic objectives then you have got that coherence. We can keep talking about multiple measures and having balancing measures, but we are not talking about excessive numbers of measures. We are talking about a coherent set that fit together and tell a story, which is important.

I think the other thing in terms of performance reporting being an onerous exercise links back to Dr Ferreira's point on the purpose of the system. People commonly think of performance measurement systems as being there for performance evaluation and judgement and assessment. To some extent that works in the private sector where you can link it to incentive payments and the like, but in the public sector it is different. It is different in the private sector as well. They are still being used for not only performance evaluation but resource allocation and, most importantly, communication within the organisation and with stakeholders. That is tremendously important here. We keep talking about communication across agencies and communication between the Committee and the different agencies. The way to do that is to have, as part of the purpose of your performance measurement system, a communication goal so that you are measuring things that will facilitate a discussion going forward rather than an evaluation process.

CHAIR: Perhaps rather than an add-on system there should be a review of the reporting requirements in order to focus it?

Ms HUMPHREYS: Absolutely, because that is the danger. If you just add to the existing system then you keep adding to the performance measures book and that is when you have got too many measures to be useful.

Professor SMITH: The risk that you run is that it is seen as an add-on. There is a lot of literature that basically supports the idea that such measures typically are not treated as seriously as they would be if they were part of a new system in their own right. I would be very much against the idea of new measures being tacked on to existing reporting requirements, partly for the burden issue and partly for the message that it sends in terms of the importance of those measures as well.

Professor CHENG: If the performance measures are perceived as something that helps to make decisions that would be information they would gather anyway. Technically, including additional reporting is not a lot of burden if these measures are actually useful. The burden happens when you report something that you do not think is very useful information, but you just want to report it. If we can look at measures that are actually helpful then it will be something that they want to gather and they will want to know about it. It is then a matter of maybe reformatting it for reporting purposes, in which case I would not think there would be a lot of extra burden.

Professor CUGANESAN: I am sure you are all expert in terms of where this is at, but there are numerous reform initiatives within the public service. I am not sure where this is at, but certainly the Schott Report made reference to changing the way that performance was designed and reported on in the New South Wales public service. Ensuring that there is some alignment between the various initiatives I think is critical, because otherwise you potentially get conflicting reform agendas and programs. That is something that I would bring to your attention, which no doubt you are already aware of.

CHAIR: Thank you. Reform fatigue is a real problem, especially if people cannot see the benefit of the reform and the process they are going through. Everyone reapplying for their jobs is very time consuming and stressful.

Professor BROWN: One of the most important things that the Committee could do, which ties in with that, would be to recognise that this is very much a work in progress, but what the Committee is doing is strategically significant not just for New South Wales and Australia but internationally. People are wrestling with this issue everywhere and no one is ahead of the game. One of the most useful things the Committee could do is to develop the right package of indicators. The Committee has to come up with its package as to what it

should look like now but it should include a recommendation to the Government that it initiate and support a project to continue the process of developing the indicators, establishing whether they are working, and to do that within New South Wales and collaboratively with other States and Territories. Transparency International would be interested in being involved in a project like that.

I am thinking back eight ten years ago when we were having these discussions. It is easy for this to keep on happening and going back to not having the right indicators and trying to invent a new package. I think it would be strategic for the Committee to recommend that the Government invest in a project which will tie it together and keep the development going. It needs to involve not just this Committee but the Public Accounts Committee in terms of how the Auditor-General is evaluated and how the Independent Commission Against Corruption committee evaluates the ICAC. There are a lot of people with a strong interest in this. That should not hold up the Committee enunciating what it thinks its package should be for the agencies you oversight, but I think it will help this Committee and successor committees if it is recognised that this is a work in progress and invested in getting more experts involved in this process.

CHAIR: When I was a 19-year-old university student doing government studies at Sydney university we were taught that a properly designed public policy has a beginning, a middle and an end, with tools for evaluation designed into the policy in the first place. It is a hopeless proposition to get people in at the end when nothing has been collected. It staggers me that nowhere in government do I see evaluation tools designed at the beginning of the program. That amazes me considering the sums of money that are being invested. Professor Prenzler, do you have any concluding comments?

Professor PRENZLER: First, congratulations on doing this. It is refreshing for those of us who care about these things—and good policing in particular—that more attention is being paid to making the guards of the guards, the oversight agencies, more accountable. I feel from our discussions that there is consensus amongst the people speaking today about the need for a package of measures. I would suggest a comprehensive package of measures. I believe it is possible to create meaningful and fairly cost-effective measures for performance outcomes for oversight agencies. At present I do not think there is any agency in Australia, including New South Wales, which has the full package. It is achievable, and I do not see any reason why such a model cannot be implemented.

CHAIR: What do you think should be the next step forward?

Professor PRENZLER: It would be great to see the Committee recommend a comprehensive package and take it to the Parliament.

CHAIR: Would the Committee have a package to put forward or would it be recommending that a package be developed?

Professor PRENZLER: I think there is one that can be taken pretty much off-the-shelf—perhaps with the Committee's adjustments. I think performance measures should be built into the governing legislation of these bodies. It is a matter of a second process where the legislation is amended organisations are required by the governing legislation to provide better reporting measures, and the core ones need to be specified.

Professor SMITH: I am not sure about the off-the-shelf performance measurement system. In keeping with the suggestions that have been made earlier today, my view would be more that the Committee recommend that a system be developed rather than saying here is the system and imposing it on folk. I think you will encounter significant behavioural resistance if you say, "Here is a system".

CHAIR: Is it a good way to frame the discussion? We have encountered going to organisations with a blank sheet of paper and all we are finding out is that nothing works and they do not need it.

Professor SMITH: I agree but I think there is hopefully middle ground between a blank sheet of paper and a fully fledged system.

CHAIR: And negotiation in the middle?

Professor SMITH: Yes. Professor Prenzler has talked about his own document which could form a starting point or work done in other jurisdiction could form a starting point but I would be reluctant to see this Committee set down what it thinks the system should be and impose it.

CHAIR: I understand that. The Committee will not be doing that, but it would love to recommend a way forward. If there is a model we can point to with potential that would help; otherwise we will not see anything happen in my lifetime. If we have experience to build on it would be great to progress that.

Mr LEE EVANS: I feel like we need to peel away the behavioural backlash. As soon as the Committee says it is looking at a new system the agencies think it is on top of what they have currently got. The Committee will probably have to set the agenda as far as reforming the system and that means everything you are currently doing will be reviewed so that all the current reporting is peeled away and the agencies only come back on the new system rather than an onus on more reporting and more boxes to tick.

CHAIR: The agencies the Committee oversights are consumers and creators of internal red tape. They are small agencies and fully aware. The Information Commissioner said she is oversighted by 77 pieces of legislation, and that is a lot for a group of 20 people to manage. There were over 30 agencies involved in the purchase of a \$240,000 off-the-shelf case management IT system. The internal red tape is an enormous problem.

Mr LEE EVANS: It is a case of measuring twice and cutting once.

CHAIR: Can I postulate one other thing: When we measure performance perhaps we need to prioritise what we are measuring. The timeliness of the answer to the correspondence is important but it is not as important as a reduction in the number of allegations of police assault.

Professor PRENZLER: The best agency that I am aware of in this area is the Northern Ireland Police Ombudsman. They have an impoverished economy yet they run the performance measures and show effectiveness in a way that most agencies in Australia cannot. They do it with a small number of researchers, perhaps three or four. Most of their staff are investigators. You do not need a lot of money to do this well, you just have to have the will to do it properly. It seems the agencies that are willing to take on the performance measures are the high achievers.

CHAIR: Because it demonstrates success. On behalf of the Committee I thank each of you for attending. This evidence has been invaluable. I hope maybe that the Committee can continue to be in contact: we might ask for your thoughts by correspondence or over the phone. The Committee especially thanks the witnesses who have travelled from interstate. We have an enormous respect for the expertise in this room. It is exciting because I do not think this expertise has been brought together in this way previously. The proceedings have been recorded and will be valuable to the New South Wales Parliament in an ongoing way.

(The witnesses withdrew)

The Committee adjourned at 1.10 p.m.