Evaluation resources for family support

- Evaluation and innovation
- Evidence-based practice & service-based evaluation
- Ethics in evaluation
- Preparing for evaluation
- Dissemination of findings

Why evaluation?

- Quality assurance – **systematic** checking of program/service meeting standards
- Were participants satisfied? Did they benefit?
- Did they benefit *because of* the program?
- How can the program be improved/refined?
- Justify requests for further support/funding
- Was implementation true to program design?

http://chronicle.com/blognetwork/lesboprof/2012/05/31/evaluation/
Types of evaluation

- "Traditional" evaluation
  - Outcome evaluation
    - Does my program help my clients?
    - Has there been any unintended outcomes?
  - Impact evaluation
    - Longer term effects of a program (though often used interchangeably with “outcome evaluation”)
  - Process evaluation
    - Is the program being implemented in the way it is intended?
    - Is it reaching the people for whom it is designed?
    - Opens up the “black box” of complex interventions
Types of evaluation

- Participatory evaluation
  - Program managers and staff treated as partners with evaluators

- Empowerment evaluation
  - Seeks to help people in communities to learn how to help themselves and to determine what services the community needs.
  - Assumes community has better knowledge than evaluators/professionals
Evaluation may lead to change...

- **Innovation in content**
  - Example – introducing discussion of use of social networking sites in relationship education program.

- **Innovation in delivery method**
  - Example – provision of web-based family dispute resolution

- **Innovation in forging new partnerships/networks**
  - Example – external or cross-agency supervision
Implementation, not just outcomes

- Increasing need to identify:
  - How effective was the implementation of the program?
  - Mediators – did anything affect process?
  - “Black box” – may reflect effective/weak implementation
### Importance of implementation

<table>
<thead>
<tr>
<th>Implementation - the “what”</th>
<th>Implementation - the “how”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective</td>
<td><strong>Effective</strong></td>
</tr>
<tr>
<td></td>
<td>Actual benefits</td>
</tr>
<tr>
<td>Not effective</td>
<td><strong>Not effective</strong></td>
</tr>
<tr>
<td></td>
<td>Inconsistent; not sustainable; poor outcomes</td>
</tr>
<tr>
<td></td>
<td>Poor outcomes</td>
</tr>
<tr>
<td></td>
<td>Poor outcomes, can be harmful</td>
</tr>
</tbody>
</table>
“Hierarchies” of evaluation - RCTs

- Different types of evidence allow stronger or weaker conclusions to be drawn.

- Randomised-controlled trials
  - Randomised – allocation to a “control” or “experimental” group
  - If no bias in allocation, more reasonable to conclude that program made a difference
  - Not without criticism, especially when applied to service provision
“Hierarchies” of evaluation - RCTs

- Problems with RCTs in service environments (also relevant to other designs)
  - Drop out of participants, especially if at different rates from the two groups
  - Unexpected differences between groups
  - Ethical issues re: withholding treatment from control group

- Randomise Me: http://www.randomiseme.org/
“Hierarchies” of evaluation – quasi-experiments

- Use of naturally occurring comparison groups
  - Participants on a waiting list
  - Offer a different intervention, e.g. a briefer version of program.

- Greater benefits to those in program may mean it is effective – but because not randomised, you can’t say it caused changes.
“Hierarchies” of evaluation – pre- and post-test

- No comparison or control group
- Measures before and after program changes
- No real conclusions can be drawn – changes might have happened anyway…
- Better than nothing…?
Service-based program evaluation

- Criticisms of evidence-based practice
  - Removed from complex real world of service delivery
  - Reliant on very restrictive definitions of “evidence”
  - Evidence may come from multiple small-scale, less sophisticated designs
  - Evidence should not be discounted just because it doesn’t fit into “hierarchy”
Figure 1. Decision tree: Who should conduct the evaluation?

Who should conduct the evaluation?

Do we have the expertise to find out if the program works?

- Yes, enough
  - Evaluation done entirely inhouse
    - Pros:
      - Existing, deeper knowledge of agency, program and systems
      - May reduce costs
      - Makes use of existing practice wisdom
      - Are already part of the agency structure and culture
      - Can draw on existing relationships with clients and stakeholders
      - Opportunity to build skills
    - Cons:
      - May not have sufficient skills to do the evaluation well
      - May need to redistribute internal staff and resources, possible impact on service delivery
      - Adds to staff workload
      - Can be difficult to be objective

- No
  - Evaluation done solely by external evaluation professional
    - Pros:
      - Increases objectivity — offers an external perspective
      - Reduced workload — less direct involvement of staff
      - Make use of greater knowledge and expertise in technical aspects of evaluation
      - Can be more efficient
      - May be less expensive than doing evaluation badly
    - Cons:
      - Can be expensive
      - Lack of knowledge and understanding of service type, agency culture
      - Lack of knowledge and understanding of the program/target group — time required to get up to speed

- Yes, some
  - Partly inhouse, partly in collaboration with external evaluator
    - Pros:
      - Introduces an outside perspective — adds a degree of objectivity
      - Increases perception of independence of the evaluation
      - Make use of inhouse knowledge and systems
      - Draws on consultant expertise and experience of other evaluations
      - Opportunity for staff to acquire new knowledge and skills
      - Reduces need to redistribute internal resources
    - Cons:
      - Can be costly
      - Lack of knowledge and understanding of program — time required to get up to speed
      - Evaluator seen as an “outsider”
Developing a culture of evaluation

- **Benefits**
  - Helps to reinforce reflective thinking
  - Helps staff to be responsive to external demands for accountability
  - Increased confidence in having evidence of a positive impact
  - Staff gain new skills and knowledge
Challenging attitudes...

- But the program is exemplary!
  - Have realistic expectations of the level of improvement that is reasonable to expect.
- But we’ll offend program staff...!
  - Remind them it is a program, not personnel evaluation. Focus discussions around “improvement”.

[Logos and text: Australian Government, Australian Institute of Family Studies]
Challenging attitudes...

- But the program will be terminated...!
  - Funders more likely to be pressured to find a replacement – thus if result is unfavourable, more likely to lead to improvements

- But evaluation drains program resources....!
  - May be true…but the alternative is that money may be spent on ineffective programs. Evaluation more likely to lead to additional support/resources
Ethical considerations

- Ethical considerations around the data you are collecting/using are needed
  - Assess what approvals are required – by organisations, ethics committees and service users and staff.
  - How do you ensure participants information is kept confidential?
  - Processes to manage potential conflicts of interest?
No ethics required?

- If original analysis of previously collected, publicly available data
- If deemed “quality assurance”
  - Do not impose risk on participants
  - Use existing organisational data
  - Analysis done in-house by someone bound by professional code of ethics
  - Do not infringe rights/reputation of carers, providers or institution
  - Do not violate confidentiality of clients
Program logic

- Or theory of change...
- Or logic model...
- Or program theory...
- Useful?

http://slashdot.org/
Program logic

- Visually represents what is going on in a program

- Two important things
  - Relationships – logical links between each stage of program logic model (if…then…)
  - Intention – a roadmap for the program
Program logic

Program Action - Logic Model

- Inputs
  - Activities
  - Participation

- Outcomes - Impact
  - Short Term
  - Medium Term
  - Long Term

Priorities
- Consider:
  - Mission
  - Vision
  - Values
  - Mandates
  - Resources
  - Local dynamics
  - Collaborators
  - Competitors

What we do
- Conduct workshops, meetings
- Deliver services
- Develop products, curriculum, resources
- Train
- Provide counseling
- Assess
- Facilitate
- Partner
- Work with media

Who we reach
- Participants
- Clients
- Agencies
- Decision-makers
- Customers
- Satisfaction

What the short term results are
- Learning
- Awareness
- Knowledge
- Attitudes
- Skills
- Opinions
- Aspirations
- Motivations

What the medium term results are
- Action
- Behavior
- Practice
- Decision-making
- Policies
- Social Action

What the ultimate impact(s) is
- Conditions
- Social
- Economic
- Civic
- Environmental

Assumptions

External Factors

Evaluation
Focus - Collect Data - Analyze and Interpret - Report

http://www.uwex.edu/ces/pdande/evaluation/evallogicmodel.html
Logical Model for Creating Achievable and Sustainable Change Modalities over Time

1. Start thinking about creating a really confusing chart.
   - Realize that you look important if it's really really confusing.
   - Realize that you look important if you're the only one who understands it.
   - Engage graphics generator
   - Engage bullshit generator

2. Put orange box here.
   - Put paper in printer
   - Build conceptual engine
     - Analyze everything you can think of in depth
     - Capture Vision
     - Let Vision Go
     - Re-capture Vision
     - Keep close watch on vision.

3. Report Outcomes
   - Check with Clients
   - Engage 1st gear
   - Roll down hill
   - Jump start engine
   - Realize that even you don't understand any of this.

4. Achieve Goals
   - Wax Eloquent
   - Wax Car
   - Finalize Results
   - Fire chimp

5. The world is a better place because they actually paid you to do this.
   - Promote/Enhance
   - Expose/Repuse

Legend:
- Movement cycles
- Try-cycles
- Predilections
- Proclivities
Program logic

**Situation:** During a county needs assessment, majority of parents reported that they were having difficulty parenting and felt stressed as a result.

**Inputs**
- Staff
- Money
- Partners
- Research

**Outputs**
- Develop parent ed curriculum
- Deliver series of interactive sessions
- Facilitate support groups

**Targeted parents attend**

**Outcomes**
- Parents increase knowledge of child dev
- Parents identify appropriate actions to take
- Parents gain skills in effective parenting practices
- Parents better understanding their own parenting style
- Parents use effective parenting practices

**External factors:**
- Improved child-parent relations
- Strong families

**Assumptions:**

Related: [http://www.uwex.edu/ces/pdande/evaluation/evallogicmodelexamples.html](http://www.uwex.edu/ces/pdande/evaluation/evallogicmodelexamples.html)
Evaluation approaches 1

- Collect new data from key informants
- Make use of internal administrative data including program data
- Use of external administrative data
- Use of existing representative datasets (e.g. Longitudinal Study of Australian Children)
- Multiple data sources allow for greater validity and also greater depth
Evaluation approaches 2

- BUT need to balance quantity of data with quality and ability to analyse
  - What resources do you have?
  - Who is conducting the evaluation? What are the skill sets available to you?
  - How much time do you have?
  - Is this a one off or an ongoing process?
  - DO YOU NEED IT ALL?
Collect new data

- Qualitative – interviews, observations, focus groups
  - Good for measuring behaviour change, greater meaning of concepts, looking for explanations
  - Observations – e.g. improving parenting practices – doesn’t rely on self-report, but very resource intensive
  - Interviews – more costly and time-intensive than focus groups, but may get greater honesty and depth
  - Focus groups – interaction can lead to high-quality responses, but group dynamics in play.
  - Analysis takes time/resources…
Collect new data

- Quantitative – surveys, questionnaires
  - Can collect a lot of information in a short time
  - Unless using an established instrument, may take much preparation, review and revision
  - Will also need cover letters, informed consent etc, etc
  - Cultural implications – check questions will not be misunderstood
**Figure 1: Hierarchy of evaluation designs and data collection methods**

**Instruments decision tree**

1. Does an instrument exist already that you can use?
   - No
   - Yes

2. Do you have the expertise to create one?
   - No
   - Yes

3. Does it need to be adapted in some way?
   - No
   - Yes

   **In what way?**
   - Wording
   - Response scales
   - Context
   - Language
   - Format & structure
   - Length

   **Consider any copyright or commercial licensing and cost issues**

   **Adapt it, test it, review and revise it**

   **Test it again**
Validity and reliability (oh no....!)

- **Validity** – does the tool measure what it is supposed to, e.g. “I thought the presenter was really nice” is not equal to behaviour change.

- **Reliability** – if used repeatedly under the same conditions, the tool gives the same result.
Internal administrative data including program data

- Useful as can have a picture across programs
- BUT – likely data is collected for very different purposes to research and evaluation - may be hard to translate across to research context
- Data quality – consistency of data being recorded
- Accessibility – privacy and confidentiality issues, physically obtaining data
- Ideally plan ahead and establish data collections and privacy processes that support both program administration and evaluation
External administrative data

- Can be used at both community and individual client levels to
  - Understand outcomes
  - Explore outcomes and monitor change
- Can be linked to client outcomes (depending on privacy issues)
- Limitations
  - Privacy and consent issues
  - Can be time intensive to access
  - Can be in complex formats that require high level statistical skills to use
Involvement of children and young people

- Have important perspectives to offer
- Balance between ethics of inclusion versus additional harm
- National Statement on Ethical Conduct in Human Research – has special section
- Examples of AIFS research that has sought view of children and young people around sensitive issues
  - Evaluation of 2006 Family Law Reforms
  - Independent Children’s Lawyer project
Dissemination

- Evaluation reports to funding bodies, management
- Publication in journals, newsletters, research blogs, social policy blogs, your own or organisation blog
- In person at conferences, seminars, network meetings
- Practice profiles, e.g.
  - Promising Practices Network (www.promisingpractices.net)
  - Blueprints for Healthy Youth Development (www.colorado.edu/cspv/blueprints/)
How do I really matter?

Becoming a “positive deviant” to make a difference

1. Count something
   * Doesn’t matter what, but should be of interest to you

2. Write (or read) something
   * Guidelines – aifs.govspace.gov.au

3. Change - be an “early adopter”
   * “As successful as medicine is, it remains replete with uncertainties and failure”.

Summary

- Evaluation resources released as part of “Focus on...Evaluation” month in November
- Also webinar by Howard Bath, short articles, etc.
- Forum on CFCA Connect – ask questions, make comments
- Thanks and good luck!