

Methods in evaluating ICDP

The first step to take in any evaluation is of course to decide which goals the interventions aim to achieve.

Conceptions:

Increased understanding about positive care giving and a more differentiated view of the child

Self-image:

Increased self-confidence as caregiver

Parent's psychological competence:

Better ability to see the interaction from the child's point of view (mentalization)

More empathy and affect consciousness

Parent's behavior:

More sensitive, positive and differentiated relationship with the child

Child's behavior:

The child's feelings are more positive and differentiated

The child's behavior is more positive and adequate for the age

These aims can be measured in different ways. Although it is important to be restrictive about the number of methods, it is also good to have data from different perspectives.

It is also important to know that ratings on different instruments are sensitive to change. It is meaningless to use an instrument if we not that it is vary hard to bring about change.

In order to evaluate the effects of ICDP, different distinctions among the methods can be made:

1. There is first the distinction between types of data collection methods. Data can be received from the person himself or herself by self-rating instruments. The person then fills out some kind of form, rating his or her attitudes, thoughts, feelings, experiences or behavior. Or the data may come from an interview, where someone asks the subject about attitudes, thoughts, feelings, experiences or behavior. Or, finally, the data may come from ratings of observed behavior, either in real time or filmed behavior.
2. Secondly, there is the question of source. Does the rating come from the child, from the parent or other family caregiver, or from a more distant caregiver such as a teacher?
3. Thirdly, about whom is the rating made? Does the parent or a teacher rate the child, or does a teacher rate a parent's behavior?

4. Fourthly, is the method used specifically made for ICDP, measuring or rating behavior that is defined as important by the ICDP principles, such as the 8 guidelines? Or is the method a more general one, used in different studies?

In the following description of methods, they have been categorized according to rating method.

Here we will mention general instruments that are well established and published scientifically and internationally. See also the list of ICDP-specific evaluation methods in the second part of this document. The best way is to combine ICDP-specific and general instruments.

Self-assessment methods

There are thousands of self-rating scales. In this context, several different types of rating scales seem to be of specific interest.

One concerns how the parent or teacher describes what he or she thinks about having been a participant in an ICDP course. Was the course adequate for the needs of the participant? This is the question of reception.

Another concern is how the conceptions about child rearing and parenthood have may changed before and after an intervention. Did the parent change opinion with regard to harsh punishment, or about how to show feeling, *in his or her own view*?

Then there are all the instruments that try to capture the individual's self-rated socio-emotional ability. We mention here a few that we have used, but there are of course many alternatives.

Interpersonal Reactivity Index (IRI; Davis, 1996) is used to measure the individual's self-perceived degree of empathy. It contains 28 items, and the results are analyzed using four subscales: Empathic care, Perspective taking, Fantasy, and Personal worry.

Toronto Alexithymia Scale (TAS-20; Taylor et al., 1999) measures the individual's self-perceived degree of alexithymia with a questionnaire containing 20 items. The scale contains three subscales: Difficulties in identifying feelings, Difficulties in describing feelings, and Externally oriented thinking. The scale has been widely used and validated in US, Canada, Netherlands, France, Germany, Italy, Korea, India and Lithuania.

Attachment Style Questionnaire (ASQ; Feeney et al., 1994) measures the individual's self-perceived attachment. The answers are analyzed using three subscales: Secure, Avoidant, and Anxious attachment styles.

Experiences in Close Relationships (ECR; Brennan et al., 1998) is an alternative to ASQ. ECR uses 36 items to self-rate two subscales: Anxiety (over not being loved and getting abandoned) and Avoidance (of close relationships)

Security scale (Kerns et al., 1996) is a method for children. They are asked how they look upon their relationship to mother and father by answering 15 questions for each parent. The scales measure type of attachment style. The child must of course be able to read.

Although self-rating instruments are easy to administer in countries where people can be supposed to read and write, they have two apparent drawbacks. One is that they probably have a large degree of social desirability in them. People tend to rate themselves as they want to see themselves, or as they think that the researcher wants them to rate, or as a protest against the rating procedure etc. Self-rating is always a communication to the person who will look at the rating form. The other problem is that the wording and translation of items is culturally influenced, and that factor must be checked in each study. Most instruments have not been validate in other than Western countries.

Interviews

Adult Attachment Interview (AAI; George, Kaplan & Main, 1985) is primarily used to get data for rating an adult person's attachment style. This interview has been widely used in different social and cultural contexts. The rating of attachment patterns is very time consuming and requires special training.

The interview can also be used to rate a person's reflective functioning (RF; Fonagy et al., 1998), which is an operationalization of Peter Fonagy's mentalization concept. The rating of RF is much easier than the rating of attachment, and it can be made on a reduced version of the interview. We are currently testing AAI and RF in more focused ways, such as a depressed person's mentalization about depressive thoughts, or a criminal person's mentalization about his or her criminal activities. It could easily be used in order to rate a parent's ability to mentalize about his or her child.

Affect consciousness interview (ACI; Monsen et al., 1996; Lech, Andersson & Holmqvist, 2008) is used to assess a person's consciousness about several categorical affects, usually interest, joy, sadness, anger, fear, guilt, and shame. This interview has been used in several studies in Sweden and Norway.

Working Model of the Child Interview (WMCI; Zeanah et al., 1986) is a structured interview, where the parent is asked to describe his or her history of experiences with one child. Like the AAI, the answers are not rated for what is said but for how it is said, and it is supposed to

catch the parent's attachment to the child. The interview is rather difficult to rate, but it gives very valuable information about the parent's conception of the child. The interview specifically catches these areas: the parent's richness of perception, openness to change, intensity of involvement, coherence, care giving sensitivity, acceptance, and view of infant difficulty.

The ratings are summed up as a description of the parent's working model of the child, where three main categories are distinguished: balanced/secure, disengaged/avoidant, and distorted/resistant. A newer version of the interview also categorizes a Disturbed pattern.

Parent Development Interview (PDI; Aber et al., 1985) is used to assess a parent's representational model of her relationship with a specific child. The interview focuses on three central domains: the parent's view of her experience in the parent-child relationship, the parent's view of the child's experience in the relationship, and the parent's overall awareness of the relationship.

Observation rating methods

Observing the child

Strange situation procedure (SSP; Ainsworth et al., 1978). This is the golden standard for measuring attachment quality in small children. The prerequisite is that the child comes with a parent to a test room. It is usually used for children between one and two years of age.

Attachment Q-sort (AQS; Waters & Dean, 1985) uses 100 cards that are sorted by the Q-sort method. It is an alternative to Strange situation. In contrast to the Strange situation, this rating can be made in the child's home. The observer is in the child's home several hours before rating the child's behavior. AQS does validly rate attachment. In a meta-analysis, 139 AQS studies including 13.835 children found positive validity, for instance that it gives similar results as the Strange situation.

An advantage in comparison with SSP is that AQS can be used for a broader age span than SSP (12-48 months).

Observing the parent

Atypical Maternal Behavior Instrument for Assessment and Classification (Ambiance; Lyons-Ruth et al., 1999) is used to rate parents' problematic behavior towards a child. It is based on video-recorded Strange situation interactions and is thus dependent on the test room.

Child-Adult Relationship Experimental Index (Care-Index; Crittenden, 1988) uses rather short video-recorded everyday interactions. It is also used with rather small children, up to two and a half year. This method has primarily been used in clinical contexts. The rating method has to be learnt from Patricia Crittenden and is rather laborious to catch.

Maternal Behaviour Q-Sort (MBQS; Pederson et al, 1994/1999) is used to rate the mother's interaction with the child by sorting 90 cards with statements about the interaction according to the Q-sort method. The intention is to rate attachment styles from mother to child.

The method is rather easy to learn.

<http://psychology.uwo.ca/faculty/pdfs/pedmor/MaternalBehaviourQSortPage.pdf>

Observing the interaction

Emotional Availability Scales (EAS; Biringen et al., 1998) is used to rate videotaped interactions between children and caregivers. The interactions are rated on six dimensions, four of them measuring the adult's behavior and attitude (parental sensitivity, parental structuring, parental non-intrusiveness, parental non-hostility) and two measuring the child's behavior (child responsiveness and child involvement). EAS has been used in many studies in different countries. Our experience is that it is rather easy to get good interrater reliability. The ratings have good concurrent validity with similar rating instruments.

The methods need to be evaluated further in these respects:

What time do they take to administer or rate?

Are they sensitive for change?

In what respects are they culturally sensitive?

Some points about cultural sensitivity:

Looking at attachment, a large number of studies seem to have converged on the view that attachment patterns have about the same frequency in different cultures. What differs is the parent's ability to accomplish attachment security, where different cultural and social contexts seem to prescribe different parents behavior leading to secure attachment.

On many other issues, the knowledge is sparse.