European Addiction Research

Eur Addict Res 1988;4:28–41

Gabriele Lachner
Hans-Ulrich Wittchen
Axel Perkonigg
Alexandra Holly
Peter Schuster
Ursula Wunderlich
Dilek Türk
Ela Garczynski
Hildegard Pfister

Max Planck Institute of Psychiatry, Clinical Psychology and Epidemiology Unit, Munich, Germany

Structure, Content and Reliability of the Munich-Composite International Diagnostic Interview (M-CIDI) Substance Use Sections

Key Words
Reliability
Diagnostic reliability
Substance use disorders
Diagnostic instruments

Abstract
After reviewing currently available diagnostic assessment instruments for substance use disorders this paper describes the format and structure of the Munich-Composite International Diagnostic Interview (M-CIDI) substance disorder section. In addition, the test-retest reliability of diagnoses and criteria for nicotine, alcohol, illegal and prescription drugs, is reported. Findings obtained in community sample of adolescents and young adults indicate that the substance section is acceptable for almost all types of respondents, efficient in terms of time and ease of administration as well as reliable in terms of consistency of findings over time. The test-retest reliability over a period of an average of 1 month, as examined by two independent interviewers indicates good-to-excellent kappa values for all substance disorders assessed, with significant kappa values ranging between 0.55 for drug abuse and 0.83 for alcohol abuse. There was also fairly consistently high agreement for the assessment of single DSM-IV diagnostic criteria for abuse and dependence as well as the M-CIDI quantity-frequency and time-related questions. To conclude, although – unlike previous studies – this study was conducted in a community sample and not in patients and used considerably longer time intervals of more than a month between investigations, our M-CIDI reliability findings are at least as high as those from previous studies.

Introduction: The Need for Diagnostic Assessment Tools for Substance Disorders

Although there has been a great deal of national and international effort in the development of substance use control policies (for alcohol, tobacco and other drugs), the same has not been true until very recently for diagnostic and classification instruments of substance use disorders. Though a number of psychiatric instruments have been developed for assessing the symptoms, the syndrome and the impact of alcohol abuse and dependence, only very few attempts have been made until recently to construct similarly detailed tools for the assessment of drug use. At the same time there is a great need for comparable versions of instruments that can be used in different settings,

The EDSP study is supported by the German Federal Ministry for Research and Technology, Project No. 01EB94056.

KARGER
Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1022–6877/98/0042–0028$15.00/0

This article is also accessible online at:
http://BioMedNet.com/karger

Gabriele Lachner
Max Planck Institute of Psychiatry
Clinical Psychology and Epidemiology Unit
Kraepelinstrasse 2, D–80804 München (Germany)
Tel. +49 89 30622 546, Fax +49 89 30622 544

e.g. in the community, primary care, clinics, hospital settings, psychiatric facilities and special substance abuse programs for the following reasons: (1) The magnitude of the problem should be assessed by epidemiological methods in general population surveys. It is also necessary to assess ongoing changes and trends. There is, therefore, a need for instruments to gather comparable data over time. (2) Early detection is essential for the success of prevention and intervention programs for the potential or actual cases. Hence, there is a need for both efficient screening and efficient diagnostic instruments. (3) Proper assessment of cases is necessary for treatment, rehabilitation and social reintegration purposes. Thus, instruments are necessary for a better definition of patient types as well as differential assessment of substance use disorder in terms of etiology, presenting symptoms, abuse patterns, substance-related problems and other associated features. (4) Standardized instruments are necessary for comparison of symptom patterns across substances and also to delineate the course and natural history of disorders. (5) Diagnostic assessment instruments are also important in the evaluation of the intervention programs in terms of their process, outcome, cost-effectiveness, impact and acceptability.

To make a rational symptom and diagnostic assessment of substance abuse, these instruments should gather information on a large number of topics including both current and lifetime substance use, primary substance and secondary substance used, predominant route of administration, amount used (dose), duration of use, pattern of use, reason for initial use, treatment history and contact with professional or other institutions, and problems related to substance use. They should also include important demographic, socioeconomic and criminal background data (e.g. ethnicity, family status, education, employment history, family history, sources of support) so that multiple dimensions of the substance use problem and the patient as a whole can be evaluated in relation to the multiple dimensions of the substance use.

Development and standardization of assessment instruments to address these issues is a challenging task given the complexity of the problems (e.g. different substances, different diagnostic approaches, and cross-cultural diversity of substance use). Fortunately, efforts to create a 'common language' in the area of psychoactive substance use has gained momentum in the last decade.

The World Health Organization (WHO) has revised the International Classification of Disease creating its tenth version [1a], parallel to a similar revision made to the US-Diagnostic and Statistical Manual of Mental Disorders, Fourth Revision [2]. Furthermore the WHO has been carrying out a still ongoing joint project with the former Alcohol, Drug Abuse, and Mental Health Administration of USA (ADAMHA) in which scientists and clinicians all over the world have joined in a collaborative effort to better define mental disorders. Within the framework of this project two major diagnostic instruments have been developed that are cross-culturally applicable [1b, 3, 4] and that include substance use disorders [5]. With regard to the cross-cultural diversity of substance use, standardized instruments have the potential to show the similarities and differences across cultures which will certainly improve our understanding of the biological, psychological and social basis of these disorders.

A book that has recently been published by the National Institute on Drug Abuse of USA (NIDA) [6] reviews the different facets of the diagnostic assessment of substance use disorders and strongly advocates using comparable measures for evaluation. Babor [7], in his review of assessment instruments, concludes that no single instrument has yet been devised to suit the needs of different types of users and suggests gathering data into three groups: (1) variables reflecting lifetime involvement with alcohol and drugs and global problem severity; (2) developmental, etiological and substance use career variables, and (3) variables related to the severity of presenting symptoms and the nature and the pattern of substance abuse. This chapter serves as an excellent guide for a clinician or researcher who wants to create a customized assessment battery for a study. It also indicates the minimum of core variables necessary to link the data with other studies. Other helpful reviews have more recently been provided by Freyberger and Stieglitz [8] and Günthner and Stetter [9].

Diagnostic assessment forms the basis of epidemiological as well as all treatment planning and other interventions. As the problems of psychoactive substance use are complex, a multidimensional assessment is suggested to cover the use pattern, symptomatology, and consequences of the substance use. This evidently time-consuming and thus costly effort can be dealt with either within one complex multifaceted assessment instrument or within a sequential process that begins with a time-efficient 'screening', proceeding to 'assessment of problem' and ends in a 'personal in-depth assessment' if needed [10]. This paper will focus on the first of these approaches: two-stage procedures have been discussed elsewhere in more detail [11].

Currently Available Diagnostic Instruments

The advent of DSM-III and -IV and ICD-10 that use specific sets of operational diagnostic criteria within the framework of coherent classification systems has brought a new perspective to instrument development. It is now possible to translate these diagnostic criteria into questions and develop instruments for the assessment of psychoactive substance abuse. These instruments are useful for the comprehensive assessment of problems, such as level, pattern and history of use, and signs and symptoms of psychoactive substance use, but they also allow the objective derivation of diagnoses. There are different motives for the development of these instruments, especially the fully structured ones such as the Composite International Diagnostic Interview (CIDI), the Diagnostic Interview Schedule (DIS), and the CIDI-Substance Abuse Module (SAM) that are developed for application by trained non-clinicians rather than by clinicians. The following interview schedules are indicative of the new generation of diagnostic instruments developed for clinical assessment and epidemiological research.

The WHO-Composite International Diagnostic Interview and Derivations Thereof

The Composite International Diagnostic Interview (WHO-CIDI, versions 1.0 to 1.2) is a fully structured interview developed by the WHO [3]. It is a schedule primarily intended for use in epidemiological studies of mental disorders by trained nonclinicians, but it can also be used for other research and clinical purposes. The structure of the interview requires minimal judgement from the interviewer, which ensures high levels of standardization and thus comparability of results. The core version has alcohol and other drug use sections that can produce ICD-10 and DSM-III-R diagnoses. The CIDI has 15 sections in modular format and can thus be tailored according to the research questions and needs [4, 12]. The reliability and validity of the core CIDI has been established in various studies [13].

The CIDI was originally developed on the basis of the Diagnostic Interview Schedule (DIS) by the National Institute on Mental Health for the Epidemiological Catchment Area Survey [12]. A major work has been published by Helzer and Canino [14] that demonstrates what can be achieved if standardized diagnostic instruments (DIS) are used in research. The DIS has produced comparative data from 10 different cultural regions that include American Citizens including Native American and Puerto Rican populations; French- and English-speaking Canadian groups; a New Zealand population in the Pacific, and three Asian populations in Taiwan, Korea and Shanghai in China.

It also needs to be mentioned that several adaptations of CIDI have been developed. The most frequently used is the University of Michigan CIDI, that was used in a National Institute on Mental Health-funded USA national survey of prevalence and comorbidity of alcohol, drug and mental disorders [15] as well as in many more recent studies around the world. The M-CIDI, discussed in more detail below, is another such adaptation, necessary primarily, because the official WHO CIDI version 2.0 that will cover DSM-IV will not be available until the end of 1997.

Composite International Diagnostic Interview – Substance Abuse Model

The CIDI-SAM [16] is a fully-structured extended CIDI module that produces DSM-III, DSM-III-R, DSM-IV Feighner, Research Diagnostic Criteria and ICD-10 diagnoses, and more recently proposed diagnostic variables for alcohol, tobacco and nine other drugs. It was designed as an optional module to expand the substance use sections of the CIDI. The diagnostic and item reliabilities of this instrument were tested in a sample of more than 900 patients in substance abuse treatment at four sites in the USA for the field trials of DSM-IV and ICD-10 [5, 17, 18]. The work also provides information as to how the criteria compare with each other. Very briefly CIDI-SAM differs mainly from the standard CIDI in the following ways: (1) substance-specific questions on medical, psychological and social consequences are included; (2) onset and recency of use are explored for each substance; and (3) diagnostic coverage is expanded (e.g., heroin and other opiates are evaluated).

Alcohol Use Disorders and Associated Disabilities Interview Schedule

The Alcohol Use Disorders and Associated Disabilities Interview Schedule (AUDADIS) was developed by the USA National Institute on Alcohol Abuse and Alcoholism [19]. It covers both the DSM and the ICD criteria for diagnosis and provides data for individual symptoms, symptom scale scores, and the syndromal clustering of related symptoms in time. It is currently being used in a large international survey on biological markers of heavy alcohol consumption and in a longitudinal survey of substance use disorders and related psychiatric comorbidity in the USA.

Structured Clinical Interview for DSM-III-R and DSM-IV

The Structured Clinical Interview for DSM-III-R (SCID) [20] was developed by Spitzer and associates and has recently been updated for DSM-IV. Here the examiner asks the suggested probe questions but is free to ask any additional questions that he or she feels are necessary to code the indicated DSM-III symptom. It has been widely used in the USA [21–23], as well as some other countries [4, 24] with demonstrated high reliability. However, for some reason lower kappa values have been obtained for the diagnostic subgroups [4]. This might be related to the problems in reliably subtyping disorders or the differential diagnosis of affective disorders with psychotic disorders. Currently, a more detailed version for alcohol and drug use disorders SCID-A/D (Alcohol and Drug sections) is being developed in parallel with AUDADIS with both instruments sharing the same content and philosophy, facilitating their comparison at the item level.

Schedules for Clinical Assessment in Neuropsychiatry

The Schedules for Clinical Assessment in Neuropsychiatry (SCAN) are a set of instruments developed by the WHO [1b] aimed at assessing, measuring and classifying the psychopathology and behavior associated with the major psychiatric disorders of adult life. SCAN is intended for use by clinicians as a comprehensive set of schedules to collect clinicial information. The system is based on the well-known Present State Examination system with the recent addition of alcohol and other drug sections. The interview uses the clinical 'cross-examination' that explores the presence of predefined symptoms. There is an accompanying glossary that gives the definitions for each symptom and thus forms the core of the system. The SCAN system also has a syndrome checklist, the 'Item Group Checklist' as well as 'Clinical History Schedule'. For most symptoms some probes are recommended, although the interviewer is free to ask any questions to assess the presence and severity of the symptom. It is a symptom-based diagnostic tool where rules for ICD-10 and DSM-IV classifications are tailored to the computer programs (CATEGO) and hence it is relatively immune to changes in the diagnostic classification. The system includes a computer-assisted interview, computerized scoring programs and a training package [25]. It is important to note that CIDI and SCAN are comprehensive instruments covering other aspects of mental functioning and also tobacco use. To our knowledge no standardized ICD-10 and DSM-III-R tobacco dependence assessment schedules other than CIDI and SCAN exist. They may, therefore, shed light on the conceptual and methodological considerations for tobacco dependence research [26].

Checklists for Diagnostic Criteria and Symptoms

There are also ICD-10 and DSM-III-R checklists that directly assess whether the criteria sets of these systems are fulfilled. These replicate the operational diagnostic criteria as stated in the classification systems [27] and require the clinician to rate whether these are present. However, without extensive prior training the test-retest reliability of these instruments across centers is usually lower and the symptoms and patterns-of-use documentation usually lacks the necessary detail.

Common Features of Diagnostic Instruments

Several comments should be made concerning the common features of these diagnostic instruments:

(1) As Robins [28] points out, these instruments collect and record data in the smallest units possible and put all the elements of the major current diagnostic systems into a single instrument. This 'atomic' and 'multi-diagnostic' approach permits data to be analyzed for multiple purposes and may be useful following the future revisions of the classification systems.

(2) Almost all the screening and assessment instruments, except the WHO/ADAMHA instruments, have been developed into one culture, for a particular diagnostic tradition, and then introduced to other cultures. It is questionable, however, whether they really do capture all the relevant aspects of substance use as it occurs in different cultures. It is important to note that CIDI and SCAN have been developed as a result of a collaborative effort of a network of centers all over the world designed for use in different cultural settings. Translations exist in more than 15 languages and the WHO is currently undertaking a systematic study on the cross-cultural applicability of the alcohol and drug use sections of these instruments [Cotter et al., pers. commun.; to be published in Drug Alcohol Depend]. This project employs a multidisciplinary approach that integrates different methods from various social sciences (sociology, anthropology, ethnography, linguistics and psychology). This endeavor will help the standardization of instruments in a culturally meaningful way and hopefully improve our understanding of cultural specificity.

(3) Another critical issue for all these instruments is that, because of the lack of a sound database, they use different thresholds with regard to the quantity and frequency of the use of the substance. Some research has been

devoted to this problem with regard to the so-called standard drink [29], but there is little data available for other substances.

(4) With the advent of official classification systems with precise diagnostic criteria we can now make better measurements of psychopathological conditions. However, the validity of a diagnostic category, and thus the validity of diagnostic instruments still remains questionable as we lack the evidence of specific diagnostic standards in substance use disorders. Therefore, what is measured by the diagnostic instrument has to be verified not only by clinical description and delimitation from other disorders [21], but also by laboratory and family studies, and studies on course and outcome (e.g. treatment reponse). It is, therefore, important to collect data on the predictive validity of both the diagnostic concepts and the instruments.

(5) As the comprehensive assessment of an individual case is more than the derivation of a simple diagnosis, additional instruments have been developed to cover the areas of social and environmental factors related to substance use for the differential and multi-axial assessment of patients in clinical settings. These include detailed individual assessment of the case on: (1) symptoms and patterns of use including quantity and frequency; (2) problems caused by substance use; (3) familial, social and environmental influences, such as severity of the condition and disability, and other consequences of use in a more comprehensive evaluation process. An example of this type of approach is the Europ ASI (Addiction Severity Index) [review by Kokkevi and Hartgers, 30].

(6) In the context of comprehensive assessment it is also worthwhile to note that currently available laboratory markers of alcohol consumption (e.g. gamma-glutamyl transferase, mean cell volume, uric acid and high-density-lipoprotein cholesterol) are fairly insensitive, although when abnormal would certainly contribute to the diagnosis of harmful alcohol consumption. Laboratory tests for other drug use are limited basically to measuring the drug or its metabolites or both in blood or other body fluids. All they indicate is whether the person has taken that particular substance recently, with the time-frame depending on the half-life of the drug. For the long-term physical effects of the substances (e.g. on the brain, liver or other organs and systems) there is a wide range of markers to be examined to assess the harmful consequences of substance use.

(7) Finally, it should also be mentioned that comprehensive assessment instruments should also allow the evaluation of the 'neurocognitive status' of the respondent [31, 32] to ascertain the validity of the responses.

The Munich-CIDI Substance Sections – Format, Structure and Content

The remainder of this article will now be devoted to the M-CIDI sections for substance disorders. Because we have described the overall format of this instrument together with its background elsewhere in this issue [Wittchen et al., p. 18; 33, 35], this discussion will be limited to special features of the substance sections as well as its reliability.

General Features

The Munich Composite International Diagnostic Interview (M-CIDI, version 2.0) substance sections are available in a lifetime and a 12-month version in various languages [33]. Both versions should be administered using a computer, with the interviewer reading each question and subsequently entering the response or for some questions verbatim responses directly into the computer. A paper/pencil version is also available requiring, however, more training and of course subsequent entry into the computer, making paper-pencil administration more vulnerable to mistakes and errors as well as less time efficient.

The M-CIDI includes three separate sections for the assessment of substance disorders, namely for nicotine dependence, alcohol abuse and dependence as well as for drug and prescription abuse and dependence. All three modules are, consistent with the standard WHO-CIDI, fully standardized and have three parts:

Part A contains screening questions to evaluate first use, frequency and quantity of use of each type of substances; these questions are put to everybody. Unlike the standard CIDI, various questions were added to evaluate circumstances of first use, effects, and problems associated with the initiation of drug use as well as more quantitative information about the type of drug and its amount.

Part B evaluates the DSM-IV and ICD-10 diagnostic criteria for each type of substance. The evaluation of some diagnostic criteria requires several questions or the presentation of a respondent list, to make it easier for the respondent to give reliable answers. The M-CIDI diagnostic questions are administered to all subjects who acknowledge at least repeated use of the respective substance assessed. These questions are identical to those in the standard CIDI. For nicotine and alcohol as well as for some criteria in the substance section age-of-onset and recency questions were not added directly after each criteria question, but at the end of the section as part of the summary evaluation in Part C. Another special feature of

Table 1. Examples from the CIDI substance sections

Example 1: *Nicotine section* Aim is to assess criterion A(2) of DSM-IV ('withdrawal')	*M-CIDI question B16* I am going to ask you about several problems you might have had in the first few days after you quit or cut down smoking. Please look at list B1 (contains 12 withdrawal symptoms A–O). Which of these difficulties on the list did you have when you quit or cut down smoking? You can simply tell me the letters. LETTERS ARE CODED! *Question B17:* A. Did you ever start using tobacco again, to avoid or to keep from having any of these problems caused by quitting or cutting down? B. And how much did these difficulties upset you or affect your everyday life?
Example 2: *Drug section* Aim is to assess DSM-IV criterion (3)	*Question L13* A. Have there ever been times when you *used more* (insert each qualifying substance, for example cannabis) *than you intended to?* B. Or did you use (substance) for much *longer periods than you intended to?* C. And when was the first time you (repeat *italicized part* of question!)? D. And when was the last time you (repeat *italicized part* of question!)?
Example 3: *Summary question in alcohol section* Aim is to ascertain 12-month clustering of diagnostic symptoms	*Question I19* Please look at list I4. Here all the questions we just discussed are briefly summarized. I'll summarize again what you have told me; you said you (all acknowledged symptoms evaluating the diagnostic criteria are repeated)... Did any of these difficulties and things ever happen at about the same time, that is within a 12-month period? A. Which of the things from the list happened first? Which last... etc. B. And when did they happen first (code time and age) and when for the last time? (code)

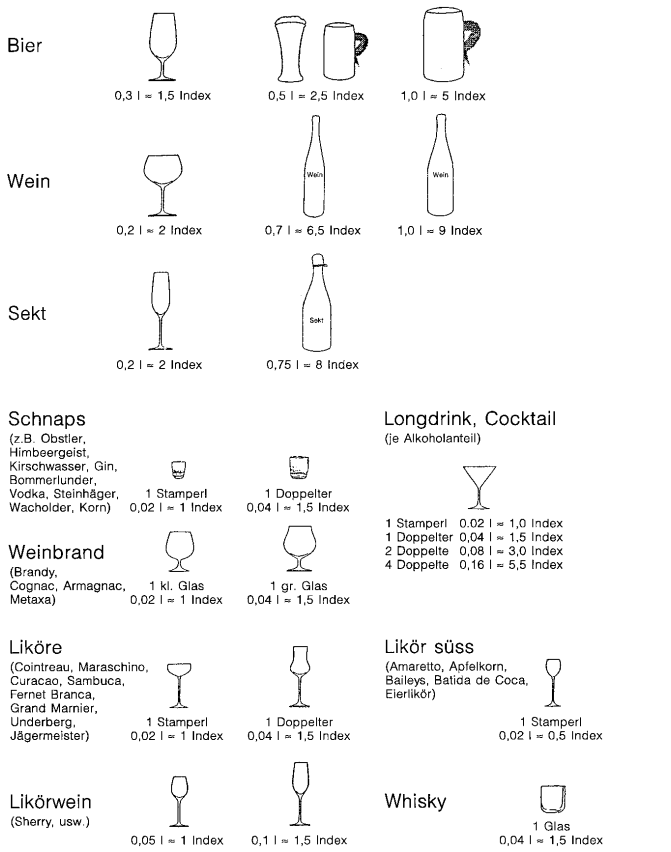
the M-CIDI is that the abuse questions are separated from the dependence questions, because some respondents find these questions offensive and might change their response pattern as a consequence. Furthermore, critical questions pertaining to problems with the police, arrests and legal consequences are asked as part of a visual aid in order to make sure that these critical questions and especially their answers cannot not be heard by any other person who might be listening to the interview.

Part C consists of a series of questions meant to assess whether the acknowledged events ever occurred together, that is in any 12-month period in the respondent's life, and, further, when the acknowledged criteria were first and last met.

Table 1 gives examples for different types of questions from each of these three parts and also clarifies how the the M-CIDI translates diagnostic criteria into specific questions.

When comparing the M-CIDI with the standard CIDI the following modifications need to be mentioned: (1) a more detailed quantity and frequency section; (2) the separation of time-related information about onset and recency into a separate series of questions at the end, and (3) the extensive use of visual aids to identify substances properly to help in the presentation of complex questions as well as to avoid asking potentially embarrassing questions out loud. Furthermore, the M-CIDI includes the option of the administration of a health belief questionnaire, assessing motivations to use, knowledge about the

Table 2. The response card with standard drink translation



risks of using the substance, the motivation for changing the patterns of use, or stopping, and subjective as well as objective barriers to potential changes.

Structure and Special Features of the Nicotine Section

This section allows for the assessment of lifetime and current use of nicotine (cigarettes, cigars, pipes, tobacco chewing or sniffing), along with information about the presence and absence of each DSM-IV and ICD-10 dependence criterion, dependence diagnosis and course over time. *Part A:* Whenever the respondent acknowledges the use of at least one of the above, 11 questions about the age and circumstances of first use, amount and various forms of reactions and effects are asked. These questions are followed by series of quantity and frequency questions along with information about onset and recency of regular use as well as questions about current and peak use for the the amount of cigarettes and their type. *Part B* is only asked to respondents who have used nicotine at least once daily over a period of at least 4 weeks, and consists of 12 symptom questions designed to evaluate the dependence criteria. Complex criteria, such as those for tolerance, withdrawal and physical as well as psychological problems related to the use are broken down into several stepwise questions or are supported by a visual aid in which the critical elements of the respective questions are represented. Whenever at least one symptom (diagnostic criterion) is acknowledged, onset and recency is evaluated in *part C* and coded accordingly into the computer. It should be noted, however, that only up to three core diagnostic criteria are dated in detail. The DSM-IV and ICD-10 clustering criterion is assessed in two alternative ways: either the respondent acknowledges having had at least three criteria within a twelve-month period, or the age of onset and recency information collected directly allows the interviewer to decide this criterion.

Structure and Special Features of the Alcohol Section

As in the nicotine section this section allows the assessment of lifetime and current use of alcohol along with information about the presence and absence of each DSM-IV and ICD-10 abuse and dependence criteria, diagnosis and course over time. *Part A:* Whenever the respondent acknowledges the use of at least one drink of alcohol, 11 questions about the age of first use, circumstances of first use, amount and various forms of reactions and effects are asked. These questions are followed by series of quantity and frequency questions along with information about onset and recency of regular use as well as questions about current and peak use for the amount and type of alcohol. In addition to the features that are identical to the nicotine section, this screen is supported by a visual aid summarizing the most prevalent types of locally available alcoholic beverages with information about their alcohol content. Throughout the subsequent assessment the M-CIDI uses the concept of the standard drink as shown in table 2 that has been adapted to local types of beverage consumption. In addition, special probe questions are used for those respondents who say they never drink any alcohol to make sure that they also take into account beer or wine consumed at a meal or festivity.

Part B is only asked to respondents who have used alcohol at least 12 times within any one year period of their life. These respondents are then asked for the exact quantity and frequency by using the standard drink definition on their response card, which can be subsequently translated by the diagnostic program into grams of alcohol per week, month, and year and over the past 12 months as well over a lifetime. Peak and current episodes of drinking behavior is also evaluated. Because it is extremely unlike-

Table 3. Summary characteristics for the drug and prescription drug section

Substances assessed	Special features
1 Cannabinoids	For each specific substance:
2 XTC and related substances	– Assessment of mode of administration
3 Other stimulants	– Frequency (one time only/2–4 times/5 and more)
4 Prescription stimulants	– Frequency and duration if 5 times or more
5 Sedatives/hypnot./anxiolytics	– First use and initial stage appraisal for all users
6 Prescribed sed./hyp./anx.	– Assessment of all dependence criteria
7 Opioids	– Assessment of age criteria were first and last met
8 Prescribed opioids	– Additional probes for frequency of criteria
9 Cocaine/crack	– Use of coded respondent lists for withdrawal and
10 PCP	– Separately for physical and psychological problems
11 Psychodelics/hallucinogens	– Use of summary dependence criteria list to
12 Inhalants/solvents	summarize
13 Other, specify	– Separate assessment of abuse criteria
	– Separate list for police, legal and criminal acts

ly that persons in our culture without a lifetime weekly consumption of 1–24 g absolute alcohol (equal to 3 standard drinks) and who, at the same time, never drink alcohol on more than two occasions per week, develop any abuse or dependence criterion, no further diagnostic questioning is pursued with these subjects.

All remaining subjects are subsequently questioned using a similar formulation to that in the nicotine section for the presence or absence of dependence criteria. In addition, response lists similar to those in the nicotine section are used. Whenever at least one dependence criterion is acknowledged, age-of-onset and recency is coded, supported by a summary statement and a response list to assess the clustering criterion. Unlike the standard WHO-CIDI, however, all eight abuse questions are asked and dated separately after the dependence portion. Possibly embarrassing questions can be assessed by use of a response list to avoid answering out loud.

Structure and Special Features of the Prescription and Illegal Drug Section

This section is designed to evaluate both symptoms and syndromes of DSM-IV and ICD-10 categories of abuse and dependence of illegal and legal psychotropic substances, such as prescription drugs. Table 3 summarizes the type of substances assessed with the variables evaluated for each of these.

The screening part of the M-CIDI drug section *(part A)* starts with the assessment of prescription drugs, either prescribed or taken on their own. Supported by a complex response list, various specific types of sedating, stimulating, or other drugs with a psychotropic potential are presented using their market names. Whenever the respondent acknowledges having used any type of these substances on more than 5 occasions either on its own, or in amounts greater or over longer periods than prescribed, the subsequent *part B* questions along with questions about quantity, frequency and onset, recency and duration are asked. Questions about the use of illegal drugs are introduced by a commitment probe, in which the person is asked whether he or she is willing to respond to these questions openly. If this is denied, the complete section is left out.

Consistent with the standard WHO-CIDI the respondent has to acknowledge having used the respective drug on at least five occasions to qualify for the subsequent quantity, frequency, onset, recency and duration questions as well as the diagnostic criteria questions. Further questions ask for mode and route of administration as well as age, circumstances, and effects of first and initial stages of the substance use. As in the alcohol section, abuse questions are grouped at the end of the dependence section in a separate module along with criterion-specific questions of onset and recency of the abuse symptoms.

The drug section is especially complex and difficult to administer by paper and pencil, if the respondent has acknowledged the use of more than one substance more than five times. The use of the computerized version shortens administration time tremendously and reduces potential errors of administration considerably. Diagnostic questions in part B are asked for every single type of substance, along with probe questions and onset and recency information.

Table 4. Design, respondents, interview and interviewer characteristics

Time interval between test and retest (n = 60)	
Average length (days): mean (SD)	38.5 (53.7)
Number (%) short interval (<4 weeks)	40 (67.8%)
Number (%) long interval (>4 weeks)	20 (32.2%)
Respondents (n = 60)	
Females	42 (70.0%)
Males	18 (30.0%)
Mean age (SD)	22.8 (8.1)
Still in school/education	24 (40.0%)
Employed	17 (28.3%)
Other (hoursewife, unemployed)	19 (31.7%)
Married	7 (11.7%)
Single	51 (85.0%)
Other	3 (3.3%)
Interview duration, min	
Nicotine (mean (SD) and *range*)	3.6 (5.1) *(0.2–19.4)*
Alcohol (mean (SD) and *range*)	7.6 (5.3) *(0.3–18.9)*
Drugs (mean (SD) and *range*)	5.1 (5.8) *(0.4–32.6)*

What Is the Reliability of the M-CIDI Sections?

Most of the M-CIDI substance questions as well as the algorithms of the M-CIDI have already been studied in various studies prior to the development of our additional modifications. However, only a few studies have examined how the CIDI works in 14- to 24-year-olds and even fewer have used community samples. Thus we decided to primarily investigate the reliability and validity in a community sample of young respondents and to focus particularly on an assessment of our modifications and additions. In this section we will first present findings from previous studies before reporting the findings of a test-retest study of 60 community subjects.

Previous Studies

A systematic review of several available test-retest and interrater reliability studies of the CIDI substance sections [13] demonstrated that substance use disorders are among the most consistent and reliable diagnostic sections covered by this instrument. Interrater reliability was found to be almost perfect in a large cross-national study (n = 575), with test-retest findings ranging between kappa values of 0.73 (illegal drugs), 0.79 for tobacco and 0.78 for alcohol use disorders [13]. Further studies also ascertained high acceptability and feasibility [5]. With these findings in mind we were now interested to see whether the M-CIDI produced similarly satisfactory high coefficients in young adults from the community.

M-CIDI Test-Retest Reliability Study

Methods. The basic *design* of this study is almost identical to previous reliability studies conducted with the CIDI [34] to allow for direct comparisons with previous findings: Respondents, after having had a preliminary CIDI interview, were invited for a second independent identical CIDI interview by a different interviewer, not knowing the findings of the first interview. Subjects were informed that the two interviews were not a memory test, but part of a methodological investigation in order to examine how well the interview worked. Further, they were informed that the retest interviewer was blind to the findings of the first interview, and thus, they should answer to the best of their knowledge and judgement in the particular assessment, not assuming that symptoms indicated in the test interview did not need to be reported again in the retest interview. Given the difficulties in recruiting cases from the community at two occassions, it was not feasible to have the short test-retest intervals of 1–3 days as in most previous studies. The mean time interval between investigations in our study was 38.5 days (SD: 53.7), ranging from 7 to a maximum of 112 days. Table 4 summarizes the average and range of time intervals between the test and the retest interviews along with information about the 60 participants in the study.

All *subjects* were sampled from respondents participating in the baseline investigation of a longitudinal 5-year study on prevalence, risk factors and incidence of mental disorders in a representative sample of 3,021 community residents in the greater Munich area, aged 14–24 years at the time of the interview [35]. After completion of the baseline interview (= test interview), each incoming interview was checked by the survey administrator for the number of symptoms in each diagnostic section. For the retest interview only subjects with at least one symptom in anyone of the M-CIDI sections were approached, in order to avoid a high proportion of definite noncases that could artificially inflate diagnostic concordance rates as well as to increase the number of cases with at least one diagnosis. To ensure sufficiently high symptomatic base rates, the survey administrator made sure in his selection of incoming consecutive interviews, that at least 5 cases were represented in each major CIDI section. In his selection process he first started with recruiting subjects with positive symptom questions in the nicotine section, and

Table 5. Diagnostic test (T)-retest (RT) reliability (base rates, kappa, percentage) of DSM-IV diagnoses of abuse and dependence of substances

DSM-IV diagnosis	Cross-tabulation			% agreement	kappa (Yules Y)
Nicotine dependence		T-n	T-y	90.0	0.64** (0.73)
	RT-n	47	2		
	RT-y	4	7		
Any alcohol disorder		T-n	T-y	93.3	0.78*** (0.82)
	RT-n	47	2		
	RT-y	2	9		
Alcohol abuse		T-n	T-y	95.0%	0.83*** (0.87)
	RT-n	48	1		
	RT-y	2	9		
Any drug disorder		T-n	T-y	95.0%	0.64*** (0.80)
	RT-n	54	1		
	RT-y	2	3		
Drug abuse		T-n	T-y	95.0%	0.55* (0.76)
	RT-n	55	1		
	RT-y	2	2		

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

subsequently added respondents with somatoform, anxiety, affective, eating disorder, alcohol, other substance disorder symptoms. Financial and personal resources only allowed the examination of 60 cases, all of whom agreed to participate.

24 *interviewers*, 16 women and 8 men, conducted at least 4 of the 120 test and retest interviews. The mean age of interviewers was 28.4 years (range 21–47 years). Fourteen of them were psychologists in postgraduate training in clinical psychology, six were students and four were professional health survey interviewers with various non-clinical backgrounds. All had significant experience in both the paper and pencil as well as the computerized version from participating in at least two separate 4-day training sessions using the M-CIDI and had completed at least five interviews in the field under the supervision of an experienced CIDI user.

Analysis. Diagnostic concordance was calculated by using the kappa statistics, additionally Yules Y coefficient and percentage agreement is indicated, along with a cross-tabulation of the raw data. Agreement for continuous variables is measured with the intra-class coefficient (ICC).

Results

Duration and Acceptance

In accordance with previous experience from the above-mentioned method studies, all respondents found the interview highly acceptable and had no complaints about length. The overall length of the complete M-CIDI is about 75 min, with the three substance sections taking together on average about 16.3 min (table 4) in our community cases. The alcohol section with an average 7.6 min was the longest. The range, however, was considerable with the longest interview taking 32.6 min for the drug section alone. The longest administration time observed for alcohol and nicotine was 18.9 and 19.4 min, respectively.

Diagnostic Agreement

Table 5 summarizes the kappa and Yules Y agreement coefficients with their significance level together with a cross-tabulation. Although base rates are low especially for drug disorders, diagnostic agreement in terms of kappa were found to be all significant and good for nicotine ($\kappa = 0.64$), acceptable for drug abuse ($\kappa = 0.55$) and excellent for alcohol use disorders. The Yules Y values, which are less sensitive to low base rates are consistently higher. The examination of disagreement reveals a slightly higher

Table 6. Test (T)-retest (RT) reliability (base rates, kappa, percentage) of DSM-IV diagnostic criteria for abuse and dependence

DSM-IV diagnostic criteria	Cross-tabulation			% agreement	kappa (Yules Y)
I. Abuse criteria					
A1: Failure in major role obligations		T-n	T-y	96.8	0.49 (0.60)
	RT-n	57	0		
	RT-y	2	1		
A2: Recurrent use in hazardous situations		T-n	T-y	93.3	0.79*** (0.83)
	RT-n	46	2		
	RT-y	2	10		
A3: Recurrent legal problems	no sufficient base rate				
A4: Continued use despite persistent social problems		T-n	T-y	98.3	0.85*** (0.85)
	RT-n	56	1		
	RT-y	0	3		
II. Dependence criteria					
1 Tolerance		T-n	T-y	85.0	0.59*** (0.65)
	RT-n	41	6		
	RT-y	3	10		
2 Withdrawal		T-n	T-y	90.0	0.69*** (0.74)
	RT-n	46	2		
	RT-y	4	8		
3 Compulsive use		T-n	T-y	86.7	0.67*** (0.71)
	RT-n	39	5		
	RT-y	3	13		
4 Unsuccessful desire to stop/cut down/control		T-n	T-y	86.7	0.69*** (0.72)
	RT-n	37	5		
	RT-y	3	15		
5 Great deal of time using and recovering from drug		T-n	T-y	88.3	0.57** (0.63)
	RT-n	49	2		
	RT-y	2	7		
6 Narrowing of repertoire		T-n	T-y	90.0	0.52** (0.67)
	RT-n	50	2		
	RT-y	4	4		
7 Continued use despite physical/psychological problems		T-n	T-y	88.3	0.68*** (0.72)
	RT-n	42	3		
	RT-y	4	11		

** $p < 0.01$; *** $p < 0.001$.

number of cases found by the retest interviewer, except for alcohol disorders.

Agreement on Diagnostic Criteria

Table 6 summarizes the agreement with regard to the assessment of each DSM-IV criterion for abuse and dependence across the three types of substances assessed. In interpreting the abuse criteria one should consider first that there is no diagnosis of abuse of nicotine and, secondly, that the base rates for criteria A3 and A1 are too low for a meaningful interpretation. The most frequently mentioned abuse criterion is the recurrent or persistent use of the substance in hazardous situations with an almost excellent kappa coefficient of 0.79 and few dis-

Table 7. Agreement of quantity-frequency questions as well as onset and recency information in the substance section; intra-class coefficients

Variable	n	ICC	MSB	MSW	F value
Number of cigarettes smoked	32	0.880	166.1	10.6	15.7
Amount of alcohol consumed (standard drink)	50	0.947	38.2	1.1	1.8
Age of first nicotine use	48	0.834	2.9	0.27	11.0
Age of first alcohol use	20	0.964	99.6	1.8	55.3
Age of first illicit drug use	10	1.00	543.1	0.10	5431.0

ICC = Intra-class coefficient; MSB = mean square between subjects; MSW = mean square within subjects.

crepancies. Abuse criterion A4 (continued use, despite persistent problems) shows only one discrepant case and a kappa of 0.79; however, the base rate of 4 cases is quite low. Among dependence criteria 'unsuccessful efforts to cut down on use' (4), compulsive use (3) and tolerance (1) were the most frequently acknowledged criteria. All kappa values are significant and range between a low of kappa of 0.52 for narrowing of repertoire and a high of kappa of 0.69 for withdrawal syndrome and cutting down. A closer exploration of substance-specific discrepancies reveals, with one exception, no interpretable differences for type of substance. However, most of the discrepancies in criterion 6 (narrowing of repertoire) can be found among nicotine users.

Agreement on Quantity and Frequency as well as Time-Related Questions

Taking the ICC as a yardstick (table 7), there is excellent agreement both in terms of quantity-frequency estimations made in the test and retest interview, as well as age-of-onset rating for substance use. In addition, agreement for the M-CIDI distinction of subjects who have just tried the substance once, repeated use (2–4 times) and regular use (more than 4 times), reveals high agreement coefficients; the resulting weighted kappa is highly significant ($\kappa = 0.72$).

Discussion

With this paper we have provided a detailed overview of the M-CIDI's substance sections format, structure and content, supplementing a more detailed presentation of the M-CIDI overall reliability in a different paper [35]. Although many, especially criteria-based questions of this instrument are identical with those in the standard CIDI, our modifications have improved the flow and the ease of administration considerably by, for example, more extensive use of response lists as well as the separation of dependence and abuse questions. The M-CIDI also offers various additions, especially tailored to research in the field of adolescent and young adult substance abuse. Among these, the module for assessing variables related to the initiation of substance use may prove to be very helpful. Compared to the administration times for the paper and pencil version, the average administration time has been significantly reduced for the substance sections by about 20% [13]. This may be due to the computerized administration as well as the more frequent use of respondents lists. Although initially we had several concerns about the use of laptop interviewers for such diagnostic purposes, no respondent investigated objected to the computerized administration with an interviewer reading the questions from the screen.

In terms of the psychometric properties we first need to acknowledge several limitations. First the base rates for the test-retest examination was fairly low, not allowing us to examine all M-CIDI questions in detail. Furthermore, the same limitation did not allow us to trace back in detail why discrepancies between the two administrations had occurred. Thirdly, we studied the reliability in a fairly young community sample. This is an advance compared to previous studies that had examined reliability primarily in identified patient samples. At the same time we need to acknowledge that we had only a few severely ill subjects with a substance disorder. Thus our findings should be generalized to community cases not to patient populations, where we might expect slightly different findings.

In summary our reliability findings obtained in a random community sample of 60 adolescents and young adults

indicate that the substance section is acceptable for almost all types of respondents, efficient in terms of time and ease of administration as well as reliable in terms of consistency of findings over time periods of at least a month. The test-retest reliability over a period of an average of one month, as examined by using two independent interviewers indicates good to excellent kappa values for all substance disorders assessed, with significant kappa values ranging between 0.55 for drug abuse and 0.83 for alcohol abuse. There was also fairly consistently high, though slightly lower agreement for the assessment of single DSM-IV diagnostic criteria for abuse and dependence. These test-retest findings compare well with, or are even better than, most previous studies using a similar design with shorter time intervals [13], both with regard to diagnosis as well as to the onset and quantity-frequency questions [36].

To conclude, although this test-retest – unlike to previous studies – was conducted in a community sample and not patients and used a considerably longer time interval of more than a month between investigations, our findings demonstrate that the CIDI substance use sections produce fairly consistent diagnostic and symptom data over a period of at least a month, even if many different trained interviewers are involved.

References

1a World Health Organization: International Classification of Disease (ICD-10). Geneva, World Health Organization, 1991.

1b World Health Organization: Schedules for Clinical Assessment in Neuropsychiatry (SCAN). Geneva, World Health Organization, 1992.

2 American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, ed 4. Washington, American Psychiatric Association, 1994.

3 World Health Organization: CIDI-CORE: Composite International Diagnostic Interview, version 1.0. Geneva, World Health Organization, 1990.

4 Wittchen H-U, Robins LN, Cottler L, Sartorius N, Burke J, Regier D, participants of the Field Trials: Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview (CIDI): Results of the multicenter WHO/ADAMHA Field Trials (wave I). Br J Psychiatry 1991;195:645–653.

5 Cottler L, Robins LN, Grant B, Blaine J, Towle I, Wittchen H-U, Sartorius N, Burke J, Regier D, Helzer J, Janca A: The CIDI-core substance abuse and dependence questions: Cross cultural and nosological issues. Br J Psychiatry 1991; 159:653–658.

6 Rounsaville B, Horton AM, Sowder B (eds): Diagnostic Source Book. Rockville, National Institute on Drug Abuse ADAMHA/USA, 1992.

7 Babor T: Alcohol and drug use history, patterns and problems; in Rounsaville B, Horten AM, Sowder B (eds): Diagnostic Source Book. Rockville, National Institute on Drug Abuse, ADAMHA/USA, 1992.

8 Freyberger HJ, Stieglitz R-D: Diagnostic instruments for the assessment of disorders due to psychoactive substance use. Eur Addict Res 1996;2:24–128.

9 Günthner A, Stetter F: Rating scales in the diagnostic process of alcohol dependence and related disorders. Eur Addict Res 1996;2:129–139.

10 Committee for the Study of Treatment Rehabilitation Services for Alcoholism, Alcohol Abuse Institute of Medicine: Broadening the Base of Treatment for Alcohol Problems. Washington, National Academy Press, 1990.

11 Üstün TB, Wittchen H-U: Instruments for the assessment of substance use disorders. Curr Opin Psychiatry 1992;5:412–419.

12 Robins LN, Wing J, Wittchen H-U, Helzer JE, Babor TF, Burke J, Farmer A, Jablenski A, Pickens R, Regier DA, Sartorius N, Towle LH: The Composite International Diagnostic Interview: An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. Arch Gen Psychiatry 1988;45:1069–1077.

13 Wittchen H-U: Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. J Psychiatr Res 1994;28:57–84.

14 Helzer JE, Canino GJ (eds): Alcoholism in North America, Europe, and Asia. New York, Oxford University Press, 1992.

15 Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, Wittchen H-U, Kendler KS: Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: Results from the National Comorbidity Survey. Arch Gen Psychiatry 1994; 51:8–19.

16 Lachner G, Wittchen H-U: Das Composite International Diagnostic Interview Substance Abuse Module (CID-SAM); in Mann K, Buchkremer G (eds): Sucht. Grundlagen, Diagnostik, Therapie. Stuttgart, Fischer, pp 147–156.

17 Cottler LB, Robins LN, Helzer JE: The reliability of the CIDI-SAM: A comprehensive substance abuse interview. Br J Addict 1989;84: 801–814.

18 Cottler LB: The use of the CIDI substance abuse module (SAM) in the DSM-IV field trial. DIS Newslett 1991;8:1–13.

19 World Health Organization: Alcohol Use Disorders and Associated Disabilities Interview Schedule (AUDADIS). Geneva, World Health Organization, 1992.

20 Spitzer JBW, Williams M, Gibbon MB: First Structured Clinical Interview for DSM-III-R. Washington, American Psychiatric Press, 1990.

21 Bryant KJ, Rounsaville BJ, Babor TF: Comparison of SCID and LEAD standard diagnoses for drug and psychiatric disorders. NIDA Res Monogr 1991;105:51–594.

22 Kosten TR, Bryant K, Rounsaville BJ: The SCID: A clinical instrument for assessing psychiatric disorders. NIDA Res Monogr 1991; 105:213–219.

23 Thevos AK, Johnston AL, Latham PK, Randall CI, Adinoff B, Malcolm R: Symptoms of anxiety in inpatient alcoholics with and without DSM-III-R anxiety. Alcoholism Clin Exp Res 1991;15:102–105.

24 Skre I, Onstad S, Torgersen S, Kringlen E: Higher interrater reliability for the Structured Clinical Interview for DSM-III-R axis I (SCID I). Acta Psychiatr Scand 1991;84:167–173.

25 Wing JK, Sartorius N, Üstün TB: Diagnosis and Clinical Measurement in Psychiatry: An Instruction Manual for the SCAN System. London, Cambridge University Press, 1992.

26 Grunberg NE, Acri JB: Conceptual and methodological considerations for tobacco addition research. Br J Addict 1991;86:637–641.

27 Janca A, Üstün TB, van Drimmelen J, Dittman V: ICD-10 Symptom Checklist, version 1,0 (MNH/MND). Geneva, World Health Organization, 1992.

28 Robins LN: Assessing substance abuse and other psychiatric disorders: History of problems, state of affairs. NIDA Res Monogr 1991;105: 203–212.

29 Miller WR, Heather N, Hall W: Calculating standard drink units: International comparisons. Br J Addict 1991;86:43–47.

30 Kokkevi A, Hartgers C: Europ ASI: European adaptation of a multidimensional assessment instrument for drug and alcohol dependence. Eur Addict Res 1995;1:208–210.
31 Gillen RW, Kranzler HR, Kadden RM, Weidenman MA: Utility of a brief cognitive screening instrument in substance abuse patients: Initial investigation. J Substance Abuse Treatm 1991;8:247–251.
32 Schafer K, Butters N, Smith T, Irwin M, Brown S, Hanger P, Grant I, Schuckit M: Cognitive performance of alcoholics: A longitudinal evaluation of the role of drinking. Alcoholism Clin Exp Res 1991;15:653–660.
33 Wittchen H-U, Pfister H (eds): DIA-X-Interviews: Manual für Screening-Verfahren und Interview: (a) Interviewheft Längsschnittuntersuchung (DIA-X-Lifetime); (b) Ergänzungsheft (DIA-X-Lifetime); (c) Interviewheft Querschnittuntersuchung (DIA-X-12 Monate); (d) Ergänzungsheft (DIA-X-12 Monate); (e) PC-Programm zur Durchführung des Interviews (Längs- und Querschnittuntersuchung); (f) Auswertungsprogramm. Frankfurt, Swets & Zeitlinger, 1997.
34 Wittchen H-U: Computer scoring of CIDI diagnoses. Int J Meth Psychiatr Res 1993;3:101–107.
35 Wittchen H-U, Lachner G, Wunderlich U, Pfister H: Test-retest reliability of the computerized DSM-IV version of the Munich-Composite International Diagnostic Interview (M-CIDI). Soc Psychiatry Psychiatr Epidemiol, submitted.
36 Wittchen H-U, Burke JD, Semler G, Pfister H, von Cranach M, Zaudig M: Recall and dating reliability of psychiatric symptoms – test-retest reliability of time related symptom questions in a standardized psychiatric interview (CIDI/DIS). Arch Gen Psychiatry 1989;46:437–443.