STARS

Electronic Theses and Dissertations, 2004-2019

2005

Non-acute Cognitive Sequelae Associated With Recreational Ecstasy Use: A Meta-analysis

Tiffany Leigh Linkovich Kyle
University of Central Florida

Showcase of Text, Archives, Research & Scholarship

NON-ACUTE COGNITIVE SEQUELAE
ASSOCIATED WITH RECREATIONAL ECSTASY USE:
A META-ANALYSIS

by

TIFFANY L. LINKOVICH KYLE
M.S. University of Central Florida, 2003

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Psychology
in the College of Arts and Sciences
at the University of Central Florida
Orlando, Florida

Summer Term
2005

Major Professor: Michael E. Dunn, Ph.D.

ABSTRACT

Studies using animal models have found considerable evidence of neurological damage resulting from exposure to 3,4-methylenedioxymethamphetamine (MDMA, ecstasy). Yet, studies comparing the cognitive performance of human recreational ecstasy users to ecstasy naïve controls have produced inconsistent results. The present study is a meta-analysis of the published empirical literature on the cognitive sequelae of human recreational ecstasy use. The pooled effect size estimate for combined cognitive domains was statistically significant and moderate in size. Small to large, statistically significant aggregate effect sizes resulted for eight of the nine cognitive ability domains included in the analysis. Moderator analyses suggested that frequent ecstasy use is associated with greater cognitive impairment, cognitive impairment can occur after relatively low amounts of total lifetime cumulative use, and recovery of functioning does not occur within one year post cessation.

This work is dedicated to Kevin Kyle, Dr. Donald Penzien, and the folks at Butterfly Yoga in Jackson, MS.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS/ABBREVIATIONS

5-HIAA   serotonin metabolite

5-HT   serotonin

CI   confidence interval

ES   effect size

FDA   Food and Drug Administration

MDMA   3,4-methylenedioxymethamphetamine

INTRODUCTION

Acute administration of 3,4-methylenedioxymethamphetamine (MDMA, ecstasy) stimulates the release and inhibits the reuptake of serotonin (5-hydroxytryptamine, 5-HT). In addition to being a serotonin agonist, MDMA more subtly increases the availability of norepinephrine and dopamine. It is chemically similar to endogenous catecholamines and neurotransmitters, but it is a synthetic mescaline analogue (Buchanan & Brown, 1988). Recreational users report acute effects of euphoria, feelings of intimacy, self-confidence, and openness to new ideas (Morgan, 2000).

In the past, some psychotherapists have administered MDMA to their patients to increase empathy and self-esteem during therapy sessions (Chiarello & Cole, 1987; Grinspoon & Bakalar, 1986; Shulgin, 1986). The U.S. Drug Enforcement Agency designated MDMA as a Schedule I controlled substance in 1986, thereby ending the legal use of MDMA by psychotherapists (Steele, McCann, & Ricaurte, 1994). MDMA has been more recently considered for use in psychotherapy (Greer & Tolbert, 1998) and within the past few years, the U.S. Food and Drug Administration (FDA) has approved clinical trials of the use of MDMA as an adjunct to posttraumatic stress disorder treatment (Imperio, 2001; Mithoefer, 2003).

Despite its illegal status, ecstasy continues to be a popular recreational drug. In 2003, 4.6% of Americans responding to the National Survey on Drug Use and Health had used ecstasy at least once in their lifetimes (Office of Applied Studies, Substance Abuse and Mental Health Services Administration, 2005). Ecstasy use is increasingly becoming popular in China (Yip & Lee, 2005) and has been acknowledged as a public health threat worldwide (Kish, 2002;

Montoya et al., 2002). However, the degree to which ecstasy truly poses a public health threat

has been controversial (Green, 2004).

Reviews of animal studies note that repeated administration of MDMA has been found to

result in lasting decreases of 5-HT and 5-HIAA (serotonin metabolite) concentrations in areas of

the brain that are typically associated with attention, memory, and executive functioning in

humans (Curran, 2000; Montoya, Sorrentino, Lukas, & Price, 2002; Morgan, 2000).

Neurological effects of MDMA have been found in non-human primates for up to seven years

post drug administration (Hatzidimitriou, McCann, & Ricaurte, 1999). On the other hand,

animal studies have also found evidence of some neuronal recovery one year post exposure.

Regeneration has been found to differ from the original innervation pattern, however, with the

dorsal neocortex remaining denervated and the amygdala and the hypothalamus evidencing

recovery or hyperinnervation (Fischer, Hatzidimitriou, Wlos, Katz, Ricaurte, 1995;

Hatzidimitriou et al., 1999; Steele et al., 1994).

There appears to be a dose-response relationship between the amount of ingested MDMA

and the amount of serotonergic neuronal damage, but it is unclear how much MDMA is required

to produce long term neurological changes. Dose-dependent reductions in 5-HT, 5-HIAA,

tryptophan hydroxylase, and 5-HT uptake sites have been found in a variety of animal species

(Parrott, 2002). However, repeated lower doses of MDMA have not been found to produce signs

of neurodegeneration in many animal studies (Vollenweider, et al, 1999). One review of animal

studies concluded that long-term neurological MDMA-related damage requires either a large

single dose or several more moderate doses, taken twice daily for four consecutive days

(Morgan, 2000).   Not only are the doses required to generate lasting neurological damage in

animals contested in the literature, there is also debate about the degree to which the possible

dose-response relationship found in animals can be generalized to humans (Lyles & Cadet, 2003; Morgan, 2000; Parrott, 2002).

Furthermore, there is controversy about whether long term serotonergic damage results from human recreational ecstasy use at all, given the manner in which it is typically used by humans (Green, 2004). Although serotonergic neuronal damage has been repeatedly found in animals, the doses are high and the findings may not directly translate to humans (Parrott, 2002). In many of the animal studies, subjects were injected with large doses repeatedly over a period of days (Boot, McGregror, & Hall, 2000). MDMA is two to three times more toxic when it is injected than when it is orally administered to animals (Ricaurte, DeLanney, Irwin, & Langston, 1988). Human recreational ecstasy users do not typically inject the drug. Usually, recreational users orally consume ecstasy and many consumers use it once per week or less because tolerance for the pleasurable effects develops quickly (Morgan, 2000). Some consumers, however, use large doses to achieve the desired response if they take ecstasy more frequently (Solowij, Hall, & Lee, 1992). Although it is not clear that human recreational ecstasy users typically consume enough ecstasy to produce neuronal damage that would lead to measurable differences in long term cognitive performance (Green, 2004), there exists variability in human recreational use patterns and, consequently, there may exist variability in the associated cognitive sequelae. Some studies have found support for a relationship between cognitive performance and amount of exposure to ecstasy in humans, generally finding differences between participants variously divided into light and heavy use groups (Morgan, 1999; Gouzoulis-Mayfrank et al., 2000; Parrott, Buchanan, et al., 2002). A longitudinal study found a relationship between continued ecstasy use and cognitive performance, suggesting that the period of time that a person has used ecstasy can impact degree of deficit in functioning (Zakzanis & Young, 2001) and one study

3

identified total lifetime cumulative ecstasy use as the strongest predictor of impaired cognitive

performance (Bhattachary & Powell, 2001).

In addition to possible differences between the quantities of MDMA administered to

study animals and the amount of exposure that is typical for human recreational ecstasy users,

there are other reasons to believe that the recreational use of ecstasy by humans would be less

likely to result in neurological damage. For example, animal studies have demonstrated that use

of MDMA increases oxidative stress while marijuana use decreases it. Oxidative stress is

thought to be a major contributor to the serotonergic neurotoxicity associated with the use of

MDMA. It has, therefore, been predicted that concomitant marijuana use might reduce risk of

MDMA-related neurological damage. Many human ecstasy users also simultaneously use

marijuana and may, therefore, be less likely to experience oxidative stress (Parrott, Gouzoulis-

Mayfrank, & Rodgers, 2004).

There are, however, some reasons to believe that the recreational use of ecstasy by

humans would be more likely to result in neurological damage. For example, recreational users

often take ecstasy in nightclubs and parties where the ambient temperature is high. At the same

time, users often do not drink enough water, increasing their risk of hyperthermia. Hyperthermia

has been found to intensify MDMA-related neurological damage in rats (Boot et al., 2000).

Studies using animal models (Lyles & Cadet, 2003) have found considerable evidence for

neurological damage resulting from exposure to MDMA. Animal studies strongly suggest that

MDMA use at some dose will result in damage to serotonergic neurons (Kish, 2002). Although

the methods of studying neurological damage in humans are constrained by ethical

considerations, some researchers have found evidence of ecstasy-related damage in human

participants. Neuroimaging of human participants has found that recreational ecstasy users have

fewer serotonin transporter sites than do non-users and the amount of difference was found to be related to the degree of exposure to ecstasy (McCann, Szabo, Scheffel, Dannals, & Ricaurte, 1998). Findings of human studies, however, have not been as definitive as they have been in animal studies (Green, 2004). Some researchers studying human participants have found lower 5-HIAAA levels measured in the cerebrospinal fluid (Ricaurte, Finnegan, Irwin, & Langston, 1990; McCann, Ridenour, Shaham, & Ricaurte, 1994) while others have not (Peroutka, Pascoe, & Faull, 1987). Methodological problems associated with studies using human participants make it difficult to confirm that ecstasy use results in a chronic serotonin deficiency syndrome. In particular, many of the studies in humans rely on markers of serotonergic neuronal integrity that can be up-regulated or down-regulated independent of the number of intact neurons. In other words, the markers may not be definitive measures of neurological damage (Kish, 2002). Furthermore, adaptive changes may be made in response to damage in a brain area thereby reducing the long term results of ecstasy use on cognitive functioning (Robbins, 2005).

Numerous studies have been conducted measuring the differences in the cognitive performance between human recreational ecstasy consumers and ecstasy naïve controls. Studies have generally focused on measures of attention, verbal and nonverbal memory, general executive functioning, fluency, learning, and reaction time. These areas have been the major areas of focus in the literature because of the relationship between these domains and serotonergic pathways thought to be damaged by exposure to MDMA (Back-Madruga, Boone, & Chang, 2003; Buhot; 1997; Curran, 2000; Montoya et al., 2002; Morgan, 2000; Parrott, 2002; Reneman, Booij, Majoie, van den Brink, & den Heeten, 2001; Verdejo-Garcia et al., 2004; Verkes et al., 2001). Neuropsychological studies of functional deficits related to ecstasy use have produced inconsistent results. Some studies have found no differences in performance on

cognitive measures. Other studies have found deficits in specific cognitive areas related to serotonergic dysfunction, but not other areas and there are inconsistencies in the literature in the pattern of specific cognitive deficits (Back-Madruga et al., 2003; Parrott, 2001; Verdejo-Garcia et al., 2004). Finally, deficits in cognitive performance may be related to reversible ecstasy withdrawal symptoms. Researchers have identified a period of ecstasy withdrawal during which recreational ecstasy users experience depression, reduced ability to concentrate, irritability, and difficulty sleeping (Curran, 2000; Parrott, 2002). Some recovery of cognitive functioning has been found to be related to time since the person last used ecstasy (Bhattachary & Powell, 2001; Morgan, 1999; Thomasius et al., 2003) and recovery of cognitive functioning is frequently found following extended abstinence from many types of abused substances (Verdejo-Garcia et al., 2004). Inconsistencies in the literature on the relationship between cognitive performance and ecstasy use might be partially explained by the minimum period of abstinence from ecstasy required before cognitive testing.

It is difficult, due to legal and ethical concerns, to conduct controlled studies of the cognitive effects of repeated ecstasy use. Therefore, most researchers have studied participants who have elected to recreationally use ecstasy in the natural environment. Most recreational ecstasy users do not exclusively use ecstasy (Morgan, 1999; Parrott et al., 2000). Consequently, researchers have had difficulty disentangling the residual cognitive consequences of ecstasy use from the consequences of polydrug use. Some researchers have even found that cognitive deficits in ecstasy users were explained by marijuana use, rather than ecstasy use (Dafters, Hoshi, & Talbot, 2004). In an attempt to control for the confound, many researchers have included polysubstance users in the control group, sometimes comparing polysubstance using ecstasy consumers to polysubstance users who have not previously used ecstasy and participants

6

who do not use drugs at all. Differences in the drug use other than ecstasy between the ecstasy use groups and the ecstasy naïve control groups may partially explain the inconsistencies in the findings of studies of the long term cognitive sequelae of human ecstasy use (Parrott, 2002).

Given the popularity of ecstasy, the movement to gain FDA approval to use MDMA in psychotherapy (Green, 2004), and the uncertainty regarding the effects of use on cognitive functioning, there is a need to empirically digest the extant studies on the relationship between recreational ecstasy use and cognitive performance. Numerous review articles have summarized the research on the cognitive sequelae of ecstasy use (Back-Madruga et al., 2003; Curran, 2000; Montoya et al., 2002; Morgan, 2000; Parrott, 2000; Parrott, 2002; Parrott, et al., 2004), but the authors of narrative reviews might intentionally or unintentionally select and interpret combinations of studies in ways that support their own theoretical positions. The process of meta-analysis helps to guard against influence of the author's bias by demanding that the author find all studies testing the hypothesis. Meta-analysis also requires that the author focus on the results of the studies, as opposed to the discussion sections of the papers. Discussion sections often include conclusions that are only weakly implied by the results (Rosenthal & DiMatteo, 2001). The present study conducted a meta-analysis of the published empirical literature on the cognitive sequelae of recreational ecstasy use. A goal of the present meta-analysis was to evaluate the overall relationship between recreational ecstasy use and cognitive performance in domains related to serotonergic dysfunction. A second goal was to investigate the possible differential relationship of ecstasy use to specific cognitive domains. A third goal was to explore possible sources of the inconsistencies in findings between primary studies.

A preliminary meta-analysis of the literature was completed in 2002 (Verbaten, 2003). A total of 14 primary studies were included in the analyses and one outcome measure was

selected from each study for each of four cognitive domains (short term memory, long term memory, attention measured as percentage correct, attention measured as reaction time). It was concluded that the literature was limited and that many more studies were necessary to investigate the possible relationship between cognitive performance and recreational ecstasy use.

Some recommend that systematic reviews be considered for an update every two years (The Cochrane Collaboration, 2005). Others contend that the decision to update a meta-analysis should be based on the number of suitable new primary studies that have been published since the previous meta-analysis, as different literatures progress at different rates (Barrowman, Fang, Sampson, & Moher, 2003). The literature focusing on MDMA has increased at an exponential rate in recent years. Publication climbed from approximately 30 studies related to MDMA being published per year in the 1990s to more than 100 studies being published per year in the 2000s (Green, 2004). Not all of the studies related to MDMA that have recently flooded the literature are studies of the non-acute cognitive sequelae associated with recreational ecstasy use. However, a substantial number of studies investigating the cognitive consequences of human ecstasy use have been published since Verbaten's (2003) study was completed in 2002.

The present meta-analysis replicated, updated, and extended the previously conducted meta-analysis. The present meta-analysis updated the former analysis by including studies that were not published at the time of the previous meta-analysis. A total of 36 primary studies were included in the present analysis, more than doubling the study-level sample size of the previously conducted meta-analysis. Inclusion of additional studies increases the confidence that the pooled effect size estimates approximate the population parameters (Rosenthal, 1984). The expanded literature also permitted additional sub-analyses of cognitive domains related to serotonergic dysfunction that were not included in the previous meta-analysis. Verbaten (2003) limited the

8

analysis to measures of immediate and delayed verbal recall, reaction time, and % correct as a measure of attention. In addition to including measures of immediate and delayed verbal memory and reaction time, the present meta-analysis expands to measures of immediate and delayed nonverbal memory, executive functioning, fluency, learning, and extensive measures of attention. Many of the primary studies included in Verbaten's (2003) analysis used multiple measures of cognitive functioning. Yet, Verbaten (2003) limited inclusion to one to four cognitive outcome measures per study. Consequently, a large amount of data from the primary studies was excluded and there was an increased chance of bias in the pooled estimates. The present meta-analysis used a method of effect size aggregation that includes all relevant cognitive measures from primary studies while guarding against sample size inflation. Verbaten (2003) tested total cumulative ecstasy consumption as a potential moderating variable. No relationship was found between cumulative ecstasy consumption and effect size. Verbaten (2003) concluded that the results suggested an absence of a dose-response relationship between cognitive performance and ecstasy use. The present meta-analysis does not limit investigation of moderating potential to total lifetime cumulative use. Multiple operational definitions of amount of exposure to ecstasy were investigated. Additional moderators were also examined.

The present meta-analysis tested the following hypotheses.

1) Ecstasy use is negatively related to quality of cognitive performance, particularly as measured by assessments of attention, immediate verbal memory, delayed verbal memory, immediate nonverbal memory, delayed nonverbal memory, executive functioning, fluency, learning, and reaction time.

2) The relationship between ecstasy use and cognitive performance is moderated by the amount of ecstasy used. If a dose-response relationship exists, it would be expected that

9

larger measurable differences in cognitive performance would occur in those human recreational users with greater exposure to ecstasy than would be found in consumers with less exposure.

3)   The relationship between ecstasy use and cognitive performance is moderated by the amount of time that has passed since ecstasy was consumed.   It is hypothesized that duration of abstinence will be related to fewer differences between the ecstasy use group and the ecstasy naïve control group.

4)   The effect size representing the strength of the relationship between ecstasy use and cognitive performance will be sensitive to the patterns of drug use, other than ecstasy, engaged in by the ecstasy naïve control groups in the primary studies.

METHOD

Search strategy and sample of studies

Relevant studies were identified by database searches of PsycINFO, PsycArticles, PubMed, and Medline using Boolean combinations of the key words ecstasy, MDMA, and 3-4 methylene-dioxymethamphetamine with memory, attention, cognitive, cognition, and executive. Additionally, the reference lists of included articles and review articles were searched for studies that might be added. The search for relevant studies terminated in June of 2005. Full text of articles that met criteria for selection for further review were obtained. If there was not enough information in the title and abstract to determine if the article met criteria, the full text was retrieved. Criteria for selection for further review were:

1.   study was published in a peer reviewed journal,

2.   study investigated MDMA or ecstasy use,

3.   study was published in English,

4.   subjects were human,

5.   abstract or title suggested that a cognitive measure of participants was likely included in the study,

6.   and abstract or title suggested that the study was not limited to the relationship between cognitive functioning and acute use.

Included studies were limited to those published in a peer reviewed journal for two reasons. First, authors conducting reviews sometimes accidentally or intentionally eliminate primary studies that do not support their position (Rosenthal & DiMatteo, 2001). Limiting

11

included studies to those that can be obtained with relative ease by other meta-analysts increases the replicability of the present meta-analysis (personal communication, Donald Penzien, April 20, 2005). Increased replicability may lead to increased confidence in the conclusions derived from the analysis because readers can be more sure of how studies were excluded. Second, components of methodological quality are tested as potential moderating variables in the present analysis. However, effect sizes from primary studies were not weighted by a global methodological quality score. Decisions on how to weight the primary studies are discussed latter in this paper. While far from perfect, peer review provides some degree of methodological quality assurance.

The full text of 74 original empirical studies were obtained and reviewed for possible inclusion in the meta-analysis. Criteria for inclusion in the meta-analysis were:

1.   study used a control group that did not contain participants with a history of ecstasy use,

2.   study used behavioral measures of cognitive functioning (i.e., not limited to self-report),

3.   reported results were not limited to multivariate analyses,

4.   manuscript included enough information to compute effect size,

5.   cognitive tests measured memory, attention, and/or executive functioning,

6.   cognitive tests were not limited to measures confounded by affective content,

7.   and participant ecstasy use was prohibited on the day that cognitive testing was

        conducted.

A total of 36 manuscripts were retained for the meta-analysis. One study (von Geusau, et al., 2004) reported the results from men and from women separately. It is recommended that, if data are collected from independent groups within the same study, the outcome measures from each group can be treated as independent effect sizes from different studies (Hunter, Schmidt, &

Jackson, 1982). As data in the male group was not repeated in data from the female group, the results were treated as independent. Another manuscript (Morgan, 1998) reported the results of two separate studies using independent samples. The results of the two studies were treated as independent. However, results from the sample from one of the two studies were reported in a latter manuscript (Morgan, 1999). Results from the same sample reported in two separate manuscripts (Morgan, 1998; 1999) were combined as one study to prevent error associated with dependent effect sizes. Care was taken to investigate the possibility that studies from the same laboratory that were published close in time used the same samples. Demographics and ecstasy use patterns were compared between samples to assess for similarity. Other than Morgan (1998; 1999), no studies that met criteria for inclusion in the meta-analysis could be determined to have used the same participants. However, the possibility that some of the same participants were used in multiple studies can not be ruled out. The total number of samples treated as independent was 37.

Data from a total of 786 ecstasy consumers and 886 ecstasy naïve controls was included in the analysis. See Table 1 for the sample size for each group from each primary study. Participants were recruited from the United Kingdom (17 samples), the United States (5 samples), the Netherlands (6 samples), Germany (4 samples), Canada (2 samples), Australia (2 samples), and China (1 samples).

Participants were required to abstain from using ecstasy for a minimum of 1 to 365 days before completing cognitive measures. See Table 2 for the minimum number of days of required abstinence by study. Average cumulative lifetime ecstasy consumption spanned from 10 to 1106 tablets. Average frequency of ecstasy consumption ranged from 1.44 to 14.8 tablets per month. Average number of years participants consumed ecstasy spanned from 0.2 to 14.7 years. See

Table 2 for information about participant ecstasy use by study.   Studies varied in the degree to which the control group matched the ecstasy use group on other substance use (see Table 3) and possible moderating demographic variables (i.e., intelligence, education level, age; see Table 4). The meta-analysis included studies representing a broad range of samples in order to increase the generalizability of the findings and reduce possible bias associated with eliminating a large amount of published data (Rosenthal & DiMatteo, 2001).

Extracting effect sizes from primary studies

Effect sizes for each primary study were computed using Hedge's g. The standardized mean difference is commonly used when comparing the means from two groups, as in the present meta-analysis that compares the means of measures of cognitive performance between ecstasy consumers and ecstasy naïve controls. The standardized mean difference, however, tends to be upwardly biased in samples smaller than 20. Hedge's g is also based on the standardized difference between two means. It is equal to the difference between the means divided by the pooled standard deviations of the two samples. The result is the number of standard deviations that separates the means of the two samples. In contrast to the standardized mean difference, Hedge's g does not overestimate the difference between means in small samples (Hedges, 1981; Lipsey & Wilson, 2001). Although r is a favorite effect size measure for meta-analysts due to ease of interpretation by social scientists (Rosenthal & DiMatteo, 2001), Hedge's g is more appropriate for the difference between two groups on cognitive performance, particularly considering the common use of standardized scores in cognitive assessment and interpretation (Lezak, 1995).

14

Computed effect sizes were limited to the differences between two groups, ecstasy consumers and ecstasy naïve controls. Primary studies frequently compared multiple groups of participants, including variations of control groups (i.e, non-substance users, marijuana consumers, ecstasy consumers), multiple durations of abstinence, and multiple levels of historical quantities of ecstasy use (i.e., heavy users, light users). Effect sizes based on differences between means in meta-analysis are limited to comparisons in primary studies in which degrees of freedom are equal to one. F tests with df > 1 indicate that there is a difference, but do not indicate directionality (Hedges, 1994). Inclusion of effect sizes resulting from multiple pairwise comparisons using various combinations of the same groups leads to error due to dependence (Hedges & Olkin, 1985). Therefore, if a study included comparisons with multiple control groups (i.e., poly-substance users and non-substance users), the control group that most resembled the ecstasy group on other drug use was selected. This selection procedure was used for all studies except one. For one study (Croft et al., 2001), it was necessary to use the non-substance user control group because some members of the marijuana using control group reported that they also used ecstasy. Therefore, the marijuana using control group did not meet the inclusion criteria for the meta-analysis. If a study included multiple time points with the same participants and same measures, the first measurement session was used to minimize influence of practice effects. If a study included former users and current users, results from the former consumers were used, to minimize the contribution of drug withdrawal. If a study included groups of heavy and light users and combined group information was not included in the manuscript, information from heavy users was used to compute effect size. Heavy users were selected to maximize the chances of participants having used a sufficient amount of ecstasy to produce a noticeable difference. If a study experimentally compared the effects of an

15

intervention in ecstasy users and ecstasy naïve controls, results of cognitive measures completed before the experimental intervention were used to compute effect sizes.

Most primary studies included in the meta-analysis used multiple measures of cognitive performance. Few primary studies used the same assessment instruments or the same combinations of cognitive measures. Selection of any one measure as representative of the data from the primary studies would have been arbitrary and possibly would open an opportunity to introduce bias. Thus, with few exceptions, results from all measures from the included primary studies that assessed attention, verbal memory, nonverbal memory, executive functioning, fluency, reaction time, or learning were included in the meta-analysis. Exceptions were made for the following circumstances. Hanson & Luciana (2004) used the Affective Working Memory (Luciana, Burgund, Berman, & Hanson, 2001) measure. The measure used stimuli with affective content that has been found to produce differential performance in people who have had their brain levels of 5-HT manipulated in laboratories. The MDMA group in Hanson & Luciana's (2004) study was significantly more depressed than the control group. As the measure, Affective Working Memory, may be confounded by differences in affective functioning and differences in affective functioning is not the hypothesis currently being tested, the effect sizes from the Affective Working Memory measure were not included in the meta-analysis. Bond and colleagues (2004) used a reading measure that was similarly confounded by affective content. Only the results from the measure that used non-affective content were used in the meta-analysis. Finally, the mean for one measure in one study was likely misreported. The reported mean would result in different statistical results than what was reported in the table in the primary study manuscript. Consequently, the effect size for LGT-3 Library delayed recall from the study conducted by Gouzoloulis and colleagues (2003) was not included in the analysis.

16

Using more than one outcome effect size from a single sample within a study gives too much weight to the results of the study. Treatment of dependent outcomes as if they were derived from independent primary studies could inflate the sample size of the meta-analysis, obscure an estimate of error, and increase the chance of committing a Type I error (Wolf, 1986). Fortunately, when a study contributes more than one effect size for outcomes, they can be used individually in subgroup meta-analyses and they can be combined in a manner that does not inflate sample size to complete an investigation of the overall relationship (Rosenthal & DiMatteo, 2001).

The present meta-analysis combined data using a computer program, Comprehensive Meta-analysis, developed through funding by the National Institutes of Health by a panel of researchers with extensive knowledge of meta-analysis (i.e., Michael Borenstein, Larry Hedges, Harris Cooper, etc.). Comprehensive Meta-analysis combines multiple outcomes measured within the same sample by computing an average. This is a common method of preventing sample size inflation inherent in treating multiple outcomes from the same study as independent (Lipsey & Wilson, 2001; Rosenthal & DiMatteo, 2001). In computing an average of the variances, it is assumed that the correlation among the outcomes is zero. Measures of cognitive functioning are often correlated. Therefore, the estimated standard error provided by averaging is likely larger than the correct standard error, thereby decreasing the chance of rejecting the null hypothesis (personal communication Michael Borenstein, July 8, 2005). The greater the intercorrelations, the less chance of rejecting the null hypothesis. To prevent underestimation of effect size, computation of a composite score that weights dependent outcomes by their covariance has been recommended (Rosenthal & Rubin, 1986) and it has been found that the composite method produces larger effect sizes than does simple arithmetic averaging (Marin-

17

Martinez & Sanchez-Meca, 1999). Without access to the original data, however, it is not

possible to generate the intercorrelations between the outcomes and the practical utility of

computing the composite score is reduced (Lipsey & Wilson, 2001). Some researchers have

recommended estimating the intercorrelations to create a composite effect size that takes into

account covariance (Rosenthal & Rubin, 1986), but it is unclear if the estimations provide a more

true result. Some combinations of cognitive measures are more intercorrelated than others. The

amount of intercorrelation among outcomes within one primary study may be substantially

different than the intercorrelation among outcomes within another primary study, particularly

when primary studies include a multitude of different combinations of outcome measures. Thus,

incorrect estimates of the intercorrelations may add another source of error and increase the

chance of spuriously rejecting the null hypothesis. The present procedure takes the more

conservative approach of averaging the dependent outcomes (Rosenthal & DiMatteo, 2001). In

interpretation, it should be remembered that the effect size is likely attenuated, however, the risk

of Type I error is low.

  Effect sizes from primary studies were not assigned weights by methodological quality.

Methodological quality can affect the primary study outcome in unpredictable ways and there is

not a theoretically sound, specific weighting method that would increase the chance of attaining

a less biased estimate of the population parameter (Shadish & Haddock, 1994). Rather,

methodological quality of the primary studies was addressed in the study selection process (see

description provided earlier in the paper) and in coding of various methodological aspects for

exploration as possible moderating variables and in sensitivity analyses.

Missing data

     Most primary study manuscripts included in the meta-analysis provided enough

information to compute effect sizes for all relevant results. Only 4 included studies provided

information for computing effect sizes for significant results, but failed to provide information

for non-significant results (Curran & Verheyden, 2003; Fox et al., 2002; Reneman, Lavalaye, &

Schmand 2001; Wareing et al., 2000). Some meta-analysts have recommended assigning an

effect size equal to zero when a primary study manuscript indicates that an effect size was not

statistically significant, but provides insufficient information to compute an effect size

(Rosenthal & DiMatteo, 2001). Statistical significance, however, is based on the difference

between the group distributions and the sample size. If the sample is not sufficiently large, an

effect size that departs from zero may not be found to be statistically significant. Differences in

distributions that would be found to be statistically significant in a large sample may fail to reach

statistical significance if the sample size is small. Furthermore, an effect size with an associated

p value of .06 is more different from zero than is an effect size with an associated p value of .50,

if the sample size is the same. While problems are introduced by including only those effect

sizes for which sufficient information exists in the primary study manuscripts, additional

problems are caused by arbitrarily assigning an effect size of zero.   The present meta-analysis

does not assign an effect size of zero as a replacement for missing data. Instead, the data are

treated as missing and Rosenthal's (1979) fail safe N is used to compute the number of studies

with effect sizes equal to zero that are needed to bring p above .05.

Assessing publication bias

As nonsignificant results are less likely to be published (Rosenthal, 1979) and, when

nonsignificant results are published, authors are less likely to include information that can be

used to compute an effect size, there is an increased the possibility that any published result is a

false positive (Begg, 1994). A funnel plot was graphed to investigate the likelihood of

publication bias. The graph plots size of the effect by standard error under the assumption that

large studies are more likely to be published than are small studies, regardless of effect size and

small studies with large effect sizes are more likely to be published than are small studies with

small effect sizes (Begg & Mazumdar, 1994). Large effect sizes that result from small studies

are more likely to be due to sampling error than are large effect sizes in large samples. Larger

studies at the top of the funnel plot should tend to cluster about the aggregate effect size, if the

effect size is a reasonable estimate of the population effect size. Small studies plotted at the

bottom of the graph should show a wider spread about the aggregate effect size. A chunk

missing in one lower corner suggests publication bias (Greenhouse & Iyengar, 1994). To

provide an estimate of how confident the reader can be in the aggregated results of the included

studies, Rosenthal's (1979) fail safe N was calculated as an estimate of the number of missing

studies required to reject the null hypothesis.

Estimating the magnitude of the relationship between ecstasy use and cognitive performance

An estimate of the magnitude of the relationship between ecstasy use and general

cognitive performance was achieved by pooling the weighted effect sizes that represented the

average of the results from all cognitive measures from each primary study. The overall estimate

is the combination of many forms of cognitive assessment from participants with a variety of

ecstasy use patterns and in a variety of nations. Although some critics have spoken against

combining the results of studies that differ in a number of ways, it has been proposed that, "It is a

good thing to mix apples and oranges, particularly if one wants to generalize about fruit" (p. 68,

Rosenthal & DiMatteo, 2001). Exact replications of studies limit the generalizability of results.

Using multiple studies that measure the construct in a variety of ways increases the external

validity of the conclusions and the construct validity of the latent variable. Robust findings are

indicated when they remain constant through a variety of samples and operational definitions

(Hedges, 1994). In addition to achieving an overall estimate, sub-analyses of the following

cognitive domains were conducted: attention, immediate verbal memory, delayed verbal

memory, immediate nonverbal memory, delayed nonverbal memory, executive functioning,

fluency, reaction time, and learning. Performing multiple meta-analyses for each type of

outcome variable enhances the interpretation of the results by permitting investigation of the

differential relationships dependent on outcome domain (Kulik, 1983; Rosenthal, 1984).

Conducting nine sub-analyses would increase the chance of making a Type I error to 36.98%, if

the null is rejected at $p < .05$ for each comparison. Therefore, the results of sub-analyses are

presented with their corresponding p value and a Bonferroni adjustment was used to set the one-

tailed $Z_{critical} = -2.128$. The Bonferroni adjustment was computed using Dubey, Armatige, and

Parmar's formula for modifying Bonferroni for correlated outcomes (Sankoh, Huque, & Dubin,

1997). A one-tailed $Z_{critical}$ was selected because it was hypothesized that ecstasy users would

have more impaired cognitive performance than would ecstasy naïve controls. Readers

concerned about Type II error may refer to p values provided in the tables describing the results

and draw their own conclusions.

Both fixed effects and random effects models were used to calculate the overall effect size and the effect sizes by domain. The fixed effects model assumes that the differences in effect sizes between primary studies are attributed solely to participant-level sampling error. The random effects model assumes that there are different population effect sizes depending on characteristics of the primary studies. The differences in effect sizes are attributed to both study-level sampling error and participant-level sampling error (Hedges, 1994; Lipsey & Wilson, 2001). Fixed effects models provide more powerful tests of the overall null hypothesis, but results based on the fixed effect model cannot be generalized to studies other than those included in the sample. The random effects model permits generalization outside of the studies included in the analysis (Rosenthal & DiMatteo, 2001). The fixed effects model can be used if the true effects sizes are invariant or if their variance is predicted by identified study characteristics with no remaining unexplained variance. If true effect sizes vary as a function of unmeasured study characteristics, it is advisable to use the random effects model. The random effects model acknowledges lack of identified sources of variance by treating differences as random. However, it is still advisable to explore possible sources of systematic variance (Raudenbush, 1994; The Cochrane Collaboration, 2005). Accordingly, when significant heterogeneity exists in effect sizes, conclusions are based on the random effects model and sources of heterogeneity are investigated.

Homogeneity suggests that there is a single population parameter that is estimated by the aggregate effect size. Heterogeneity poses the potential that there are multiple population parameters dependent on study or participant characteristics. Homogeneity was evaluated using I^2 and the Q statistic with a chi-square test of significance. Degrees of freedom for the Q statistic are k -1, where k is the number of primary studies included in the analysis. As the significance

of the Q statistic is affected by sample size, other methods of heterogeneity investigation were also used. Heterogeneity was also assessed by comparing the aggregate effect size estimates resulting from the application of the fixed versus random effects models and by inspecting the amount of overlap in confidence intervals of primary study effect sizes. If there is little overlap in the confidence intervals for the results of the individual primary studies, the results are likely heterogeneous (Hedges & Olkin, 1985; Higgins & Thompson, 2002; The Cochrane Collaboration, 2005).

Investigating potential sources of heterogeneity

Tests of heterogeneity examine the likelihood that differences between effect sizes are due to sampling error or that there is likely some systematic variance among effect sizes that may be attributed to moderator variables (Hedges, 1994). Variability among effect sizes indicates that influences of possible moderating variables should be investigated. When searching for moderators, Rosenthal and DiMatteo (2001) recommend using an explorative, rather than a confirmative approach, to contribute to theory development and identify areas for further study. An exploratory process was used in sensitivity analyses, but a priori identified potential moderators were investigated (see hypotheses listed earlier in the paper). As significant heterogeneity exists in the primary studies, the method of unrestricted maximum likelihood for mixed effects regression was used to test possible moderating variables identified as continuous variables (Raudenbush, 1994). Sensitivity analyses were performed to test the potential that the resultant overall effect size was dependent on the study inclusion criteria and to explore possible sources of heterogeneity. As a portion of the sensitivity analysis, aggregated effect sizes were

repeatedly calculated with each study removed once per calculation. This was done to determine

if inclusion of any one study had a large impact on the aggregate effect size estimate

(Greenhouse & Iyengar, 1994).

RESULTS

Analysis of combined cognitive domains

An overall effect size for all included cognitive domains combined across the sample of 37 studies was estimated using the fixed effects model, ES = -0.50, CI (-0.60 to -0.40), Z = -6.22, p = .000 (see Table 5).   The effect sizes for most primary studies were in the predicted direction; however, inspection of the Forrest plot (see Figure 1) revealed that there was variance in the magnitude of the effect and the 95% confidence intervals for a number of the primary study effect sizes crossed zero. There was statistically significant heterogeneity across study effect sizes, I^2 = 57.27, Q (36) = 84.25, p = .000 (see Table 6). Therefore, the overall effect size using the random effects model was inspected.   It was statistically significant, ES = -0.52, CI (-0.60 to -0.40), Z = -6.22, p = .000, and was similar to the effect size based on the fixed effects model (see Table 5). The overall effect size using the random effects model is a moderate effect size and roughly corresponds to a correlation of 0.3 (Cohen, 1988).

Assessing the potential for publication bias

Inspection of the funnel plot (see Figure 2) for the present meta-analysis suggested that there is a degree of publication bias in the estimated aggregate effect size.   If no publication bias existed, a larger number of small studies that found no relationship between ecstasy use and cognitive performance would have been published and available for inclusion in the meta-analysis. As publication bias was suggested by the funnel plot, Rosenthal's (1979) fail safe *N*

was computed. The fail safe *N* for the overall effect size was 850 studies. The fail safe *N*

indicates the number of missing studies with an effect size of zero that would need to be included

in the meta-analysis to bring *p* above .05.

Sub-analyses for each cognitive domain

Sub-analyses for each cognitive domain were conducted (see Table 7). All aggregate

effect sizes for cognitive domains were in the predicted direction, with the ecstasy use group

performing worse than the ecstasy naïve control group. Significant aggregate effect sizes for

attention, immediate verbal memory, delayed verbal memory, immediate nonverbal memory,

delayed nonverbal memory, executive functioning, fluency, and reaction time were found,

$Z_{observed} < -2.128$. The aggregate effect size for learning was not statistically significant, $Z_{observed}$

$= -1.349$. Comparison of *Z* and *p* values among sub-analyses is less informative than is

comparison of effect size values, as *Z* and *p* values are influenced by sample size and the number

of studies included in each sub-analysis differs. Comparison of aggregate effect sizes suggested

that the strongest relationships between ecstasy use and cognitive performance were found in

measures of delayed verbal memory, ES = -0.88, CI (-1.50 to -0.26), *Z* = -2.77, *p* = .006,

immediate verbal memory, ES = -0.63, CI (-0.85 to -0.41), *Z* = -5.65, *p* = .000, and delayed

nonverbal memory, ES = -0.63, CI (-1.15 to -0.11), *Z* = -2.39, *p* = .000.

Heterogeneity

Heterogeneity for the overall effect size for combined cognitive domains, the sub-

analyses of each cognitive domain, and the sensitivity analyses was evaluated using I^2, the *Q*

26

statistic with a chi-square test of significance. I^2 is the percentage of the variability in the effect

size estimates that is attributed to heterogeneity as opposed to participant-level sampling error

(Higgins & Thompson, 2002). Heterogeneity statistics can be found in Table 6. Significant

heterogeneity was found in the overall effect size for combined cognitive domains and in most of

the sub-analyses for each cognitive domain, with the exception of executive functioning, $I^2 =$

21.84, Q (12) = 15.35, $p = .223$, and learning, $I^2 = 0.00$, Q (7) = 4.55, $p = .714$. It should be

remembered that the significance of the chi-square test for Q depends upon sample size. It can

result in highly significant values despite little variation among effect sizes when the sample size

is large (Rosenthal & DiMatteo, 2001; The Cochrane Collaboration, 2005). Inspection of

Forrest plots and effect sizes for each primary study provides further indications of heterogeneity

and can sometimes suggest sources of heterogeneity. A Forrest plot for the analysis of the

overall effect size for combined cognitive domains is presented in Figure 1. Forrest plots for

sub-analyses and sensitivity analyses are presented in Figures 3 through 15. Inspection of the

Forrest plots and primary study effect sizes suggested that effect sizes for each cognitive domain

were generally in the predicted direction. This suggested that most of heterogeneity was due to

differences in the magnitude of the effect sizes, as opposed to differences in the direction of the

effect sizes. However, there were infrequent disparate findings and most of the 95% confidence

intervals for effect sizes crossed zero.

Investigation of potential moderating variables

Heterogeneity suggests the influence of at least one moderator. A priori hypothesized

moderators were tested. As there was uneven reporting of moderator variable information in the

primary studies (see Table 2), separate meta-regression analyses were conducted for each potential moderator variable. A Bonferroni adjusted $Z_{critical}$ of +/- 2.569 was used to protect against experiment-wise error. A two-tailed test was selected since it was anticipated that some moderators would be negatively related and some would be positively related to effect size.

Amount of ecstasy consumed was measured by three variables: self-reported total lifetime cumulative ecstasy use, self-reported frequency of ecstasy use per month, and self-reported total number of years ecstasy was used. The three measures of amount of ecstasy consumed and the duration of abstinence from ecstasy before cognitive assessment were tested in four separate mixed effects regression analyses using the method of unrestricted maximum likelihood (see Table 8). Results failed to support a relationship between total lifetime cumulative ecstasy use and effect size, $Q_{model} = 0.17$, $Q_{total} = 30.39$, $Z = -0.42$, $p = 0.677$, (see Figure 16). Seven studies were not included in the meta-regression because average total lifetime cumulative ecstasy use was not reported in the manuscripts. Results supported the relationship between reported frequency of ecstasy use per month and effect size, slope point estimate = -0.09, $Q_{model} = 33.60$, $Q_{total} = 61.04$, $Z = -5.80$, $p = 0.0000$ (see Figure 17). Seventeen studies were not included in the meta-regression because frequency of ecstasy use was not reported in the manuscripts. Results failed to support a relationship between number of years ecstasy was used and effect size, $Q_{model} = 3.22$, $Q_{total} = 28.95$, $Z = 1.79$, $p = 0.073$, (see Figure 18). Ten studies were not included in the meta-regression because information about the required duration of abstinence was not quantified and reported in the manuscripts. The Cochrane Collaboration (2005) recommends conducting sensitivity analyses of meta-regressions by testing with and without extreme scores. As an extreme score may have influenced the results of the meta-regression, data from Zakzanis and colleagues' (2002) study was removed and the

analysis was repeated. Results continued to fail to support a relationship between number of

years ecstasy was used and effect size, $Q_{model} = 2.11$, $Q_{total} = 27.98$, $Z = 1.45$, $p = 0.146$, (see

Figure 19). Finally, results failed to support a relationship between duration of required

abstinence from ecstasy and effect size, $Q_{model} = 0.33$, $Q_{total} = 35.59$, $Z = -0.57$, $p = 0.567$, (see

Figure 20). Two studies were not included in the meta-regression because information about the

required duration of abstinence was not reported in the manuscripts. Although a substantial

proportion of the heterogeneity could be attributed to a potential dose-response relationship (i.e.,

frequency of ecstasy use per month), a large amount of heterogeneity remained (see Table 8).

Sensitivity analyses

Sensitivity analyses were conducted to investigate the potential influence of study

inclusion criteria on aggregate effect size estimates. The sensitivity analyses were exploratory.

Accordingly, no adjustment was made to $Z_{critical}$ to control for experiment-wise error.

Primary studies restricted participant inclusion criteria for control groups in a variety of

ways. Sensitivity analyses in the form of sub-analyses divided by other substance use of control

group were conducted (see Table 9). As control group substance use increased, aggregate effect

size for the relationship between ecstasy use and cognitive performance approached nil, but

remained in the predicted direction. The aggregate effect sizes were statistically significant for

all but the sub-analysis of primary studies exclusively using polysubstance users in the control

groups, ES = -0.25, CI (-0.573 to 0.074), $Z = -1.511$, $p = .131$. Results for the sub-analysis of

the primary studies using polysubstance consumer control groups were the same for the fixed

and random effects models due to homogeneity of primary study effect sizes (see graphic display of fixed and random effects point estimates and confidence intervals in Figure 15).

Heterogeneity statistics suggested that differences in primary study effect sizes for the sub-analysis limited to control groups consisting of polysubstance consumers could be explained by participant-level sampling error, $I^2 = 0.00$, $Q(3) = 0.07$, $p = .996$. Heterogeneity of primary study effect sizes for the sub-analysis limited to control groups consisting of marijuana consumers could also likely be explained by participant-level sampling error, $I^2 = 0.00$, $Q(5) = 1.89$, $p = .864$. To a lesser degree, support was found for the homogeneity of primary study effect sizes for the meta-analysis limited to control groups consisting of non-substance users, $I^2 = 42.96$, $Q(4) = 7.01$, $p = .135$. The homogeneity of primary study effect sizes was not supported for the sub-analysis limited to control groups consisting of a mixture of non-substance users and participants who used a variety of substances, $I^2 = 57.30$, $Q(17) = 39.81$, $p = .001$ (see Table 6). This suggests that study-level sampling error may influence the aggregate effect size for this sub-analysis. When study-level sampling error is likely, it is advisable to interpret results from the random effects model. The results from the random effects model will resemble the results from the fixed effects model to the degree that homogeneity is present (The Cochrane Collaboration, 2005). Accordingly, the effect sizes for the random effects model are presented in Table 9 for comparison among sub-analyses. Heterogeneity as well as graphic displays of point estimates and confidence intervals for both fixed and random effects models can be found in the Forrest plots in Figures 12 through 15.

Only 3 primary studies reported that the differences between the ecstasy use groups and the ecstasy naïve control groups on alcohol, marijuana, and other drug use were not statistically significant. The remainder of the studies tested the statistical significance of the differences

between the groups on fewer than all three variables, reported that one group used more than the

other, or failed to provide information about the substance use of the control group. In 19 of the

primary studies, the ecstasy use group used more marijuana, alcohol, or other drugs than did the

ecstasy naïve control group (see Table 3).

A sensitivity analysis was conducted using only studies with control groups known to be

balanced for alcohol, marijuana, and other drug use. The analysis was conducted because

independent studies have found that alcohol, marijuana, and other drug use can influence the

results of cognitive assessment (Verdejo-Garcia, et al., 2004). The resultant aggregate effect size

estimate was in the predicted direction, but was not of sufficient magnitude to reach statistical

significance, ES = -0.34, CI (-0.763 to 0.087), Z = -1.558, p = .119 (see Table 9). Results were

the same for the fixed and random effects models, due to homogeneity of primary study effect

sizes, I^2 = 0.00, Q (3) = 0.96, p = .620 (see Table 6). Graphic displays of point estimates and

confidence intervals for both fixed and random effects models can be found in the Forrest plot in

Figure 21. Note that two out of the three studies included in the sensitivity analysis were from

the same laboratory and were completed within approximately the same timeframe. Results of

this sub-analysis should be interpreted with caution due to the small number of studies included

in the analysis and the dependence in effect sizes between the two studies completed close in

time and space.

In 11 primary studies, differences between the ecstasy use groups and the ecstasy naïve

control groups on general intelligence, education, and age were reported to be found to not be

statistically significant (see Table 4). As general intelligence, education, and age can influence

the results of cognitive assessment (Lezak, 1995), a sensitivity analysis was conducted using

only studies with control groups known to be balanced for these variables. When the control

groups were balanced for general intelligence, education, and age, the effect size computed under the fixed effects model remained statistically significant and was in the predicted direction, ES = -0.42, CI (-0.577 to -0.254), Z = -5.036, p = .000. A significant amount of heterogeneity was found in the effect sizes of the included primary studies, I^2 = 73.99, Q (11) = 38.44, p = .000 (see Table 6 and Figure 22). Consequently, results of the random effects model were interpreted. The aggregated effect size computed under the random effects model remained statistically significant and was in the predicted direction, ES = -0.35, CI (-0.678 to -0.017), Z = -2.061, p = .039.

To test the sensitivity of the combined effect size estimate to the influence of individual primary studies, the overall effect size for combined cognitive domains was calculated with one study removed per calculation (Greenhouse & Iyengar, 1994). As would be expected, given the sample size of the study (ecstasy use group n = 100, control group n = 100), Yip & Lee (2005) had the largest influence on effect size. With Yip & Lee (2005) removed, the aggregate effect size remained statistically significant and continued to be moderate in magnitude, but was slightly reduced, ES = -0.47, CI (-0.61 to -0.32), Z = -6.20, p = .000, see Table 10. A Forrest plot of the one study removed sensitivity analysis is presented in Figure 23.

One study removed sensitivity analyses were also conducted for sub-analyses of each cognitive domain. Results can be found in Tables 11 through 19. Removing any one study did not greatly affect the results of sub-analyses for measures of attention, immediate verbal memory, immediate nonverbal memory, fluency, or reaction time. However, influences of single studies were evident within other cognitive domains. Removing Yip & Lee (2005) from the aggregate estimate of the effect size for measures of delayed verbal memory did not change the significance of the aggregate effect size, but substantially reduced the magnitude of the estimate,

ES = -0.58, CI (-0.85 to -0.30), Z = -4.09, p = .000 (see Table 13).   Removing McCann and colleagues' (1999) study from the aggregate effect size for measures of delayed nonverbal memory and removing Gouzoulis-Mayfrank and colleagues' (2000) study from the aggregate effect size for measures of executive functioning increased the p values above .05 and the Z values above the Bonferroni adjusted $Z_{critical}$ value of -2.128 (see Tables 15 through 16). Removing Back-Madruga and colleagues' (2003) study from the aggregate effect size for measures of learning decreased the p value to below .05, but the Z value remained above the Bonferroni adjusted $Z_{critical}$ value of -2.128 (see Table 19).

DISCUSSION

The present study used meta-analysis to combine the results of the empirical literature on the non-acute cognitive sequelae of recreational ecstasy use. Meta-analysis provides a powerful test of the relationship between two variables. Similar results repeated over many studies, even if not statistically significant within the primary studies, provide stronger evidence of a relationship than does one study with a statistically significant result (Rosenthal & DiMatteo, 2001). The resultant overall effect size for combined cognitive measures suggests a moderately strong relationship between non-acute recreational ecstasy use and cognitive performance. The overall effect size of -0.52 CI (-0.60 to -0.40) indicates that participants in the ecstasy use groups performed, on average, a little over one-half of a standard deviation below the performance of ecstasy naïve participants on combined measures of cognitive functioning. In addition, small to large, statistically significant aggregate effect sizes resulted for eight of nine cognitive ability domains included in the analysis. The one exception was the aggregate effect size for learning. Although the aggregate effect size for measures of learning was in the predicted direction, the 95% confidence interval crossed zero, indicating that the null hypothesis could not be rejected. Results of the present meta-analysis are consistent with animal studies that show serotonergic neurotoxicity as a result of exposure to MDMA (Steele et al., 1994; Lyles & Cadet, 2003). They are also consistent with results of functional neuroimaging studies that locate lesions associated with ecstasy use in areas of the brain that are involved in attention, memory, and executive functioning (Hurley, Reneman, & Taber, 2002). Finally, they are consistent with studies that correlate ecstasy-related serotonin deficits with cognitive functioning (McCann et al., 1994).

34

Effect sizes varied as a function of reported frequency of ecstasy use per month, but were not moderated by other indications of ecstasy exposure (i.e., total lifetime cumulative ecstasy use, years of ecstasy use). Effect sizes were also not moderated by duration of abstinence before cognitive testing. This is inconsistent with some empirical findings that have previously been published (Steele et al., 1994; Thomasius et al., 2003). However, a recently published longitudinal study comparing the cognitive performance of former ecstasy users to continuing ecstasy users also found no recovery of functioning in the abstinent ecstasy users (Gouzoulis-Mayfrank et al., 2005). That no recovery of functioning was evident suggested that the cognitive sequelae are not limited to possible ecstasy withdrawal symptoms. Only a small number of studies were conducted in which a minimum of one year of abstinence was required before testing. It is possible that that the relationship between duration of abstinence and cognitive functioning would be found if more studies were conducted with one year or longer required abstinence periods.

Findings were complicated by differential effect sizes identified in sensitivity analyses when various study inclusion criteria were used. Inspection of the effect size estimates based on subgroups of studies divided by type of control group (see Table 9) suggested that the relative cognitive impairment found in ecstasy users was confounded by the amount of other drugs used by the control group. As substance use (other than ecstasy) of control group increased to match substance use of the ecstasy use group, the effect size for the relationship between ecstasy use and cognitive performance approached the nil (see Table 9). It did not reach zero, but became statistically non-significant. Care should be taken in comparing the *p* values in the sensitivity analyses, as *p* is strongly related to sample size and there were differences in the number of studies that used the various types of control groups. The sensitivity analysis that included

control groups consisting exclusively of polysubstance users involved aggregating the results of only four primary studies. Consequently, conclusions based on the non-significance of the finding should be made with extreme caution. The lower the number of primary studies included in a meta-analysis, the less powerful the test of the null hypothesis and there is a greater chance of committing a Type II error (Rosenthal, 1984). Additional studies are needed to verify the results of this sub-analysis. If the null hypothesis was correctly rejected, it does not rule out a relationship between ecstasy use and impaired cognitive performance. Non-acute other substance use has been found to impair cognitive performance (Verdejo-Garcia, et al., 2004). The hypothesis tested in the present study was not that ecstasy use would be related to more severe cognitive impairment than other drug use, but that it would be related to impaired cognitive performance. That the effect size remained in the predicted direction is notable.

When the control groups included only marijuana users, there continued to be a statistically significant effect size in the predicted direction. This suggested that impaired cognitive performance associated with ecstasy use could not be completely explained by concomitant marijuana use. The effect size, however, was attenuated. Some might conclude that the attenuated effect size is evidence that concurrent marijuana use might reduce the risk of neurotoxic ecstasy effects (Parrot, et al., 2004). It is more likely that the control group was experiencing cognitive impairment related to marijuana use and, therefore, performed more like the ecstasy use group. While the evidence of sustained cognitive sequelae of marijuana use is inconsistent at best (Grant, Gonzalez, Carey, Natarajan, & Wolfson, 2003), there is evidence that chronic marijuana consumers experience associated cognitive impairment during a period of withdrawal (Pope, Gruber, Hudson, Huestis, & Yurgelun-Todd, 2001).

The strongest effect size was found in comparisons between ecstasy users and non-drug users. Using non-drug users as the control group introduces the confound that other drug use may be responsible for the difference between the two groups. As ecstasy users seldom exclusively use ecstasy (e.g., Morgan, 1999; Parrott et al., 2000), it is more appropriate to use a control group that is matched with the ecstasy use group on other substance use. Although 5 primary studies used participants without experience using other drugs in the control groups, only 3 studies demonstrated that the other drug use of the control group was not statistically significantly different from the other drug use of the ecstasy use group (see Table 3). A sensitivity analysis including only the studies that were balanced on alcohol, marijuana, and other drug use resulted in an effect size in the predicted direction, but failed to reach statistical significance. Again, the results of the sub-analysis should be viewed with extreme caution as they were based on the aggregation of 3 primary studies. Additional caution is warranted given that 2 of the 3 studies were completed in the same laboratory at around the same timeframe.

Findings of sensitivity analyses highlight the importance of cautious inclusion and exclusion criteria for participants in observational studies of the consequences associated with recreational ecstasy use. To reduce the potential confound of cognitive impairment arising from other substance use, future researchers may wish to focus data collection on participants with less severe other substance use and on a control group that is carefully matched on other substance use to the ecstasy use group. In addition, if there is substantial other drug use in the control group, researchers may wish to increase their sample size because the resultant effect size will likely be attenuated. Findings of the present meta-analysis suggest that, although it is likely that the magnitude of the effect size will change as a result of degree of inclusion of participants with other substance use, the direction of the effect size is likely to remain the same.

37

The present meta-analysis is a replication and extension of a meta-analysis of the long term cognitive sequelae of ecstasy use completed in 2002 (Verbaten, 2003). The present meta-analysis improved upon the previous analysis by including more of the outcome data from the studies that were used in the previously conducted meta-analysis. Verbaten (2003) selected one to four outcomes per study. The present meta-analysis included all relevant outcomes, permitting greater confidence that selection of which outcome to include did not bias the results. The present meta-analysis also included studies that were not published at the time of the previous meta-analysis, thereby enabling additional sub-analyses within the present meta-analysis of cognitive domains that were not included in the previous meta-analysis. Inclusion of more studies also permits greater confidence in the results of the meta-analysis. Whereas the highest fail safe N is Verbaten's (2003) study was 14, the fail safe N in the present meta-analysis was 850. Results of the present meta-analysis are similar to what was found in the former meta-analysis. In the present meta-analysis, the aggregate effect sizes for immediate verbal memory and delayed verbal memory were found to be -0.63 and -0.88 respectively. Verbaten (2003) estimated the effect sizes for immediate verbal memory and delayed verbal memory to be -1.15 and -1.36 respectively. Significant heterogeneity was found in the effect sizes from the primary studies included in both the previous and present meta-analysis. Heterogeneity suggests the presence of study-level sampling error. When the potential for sampling error exists, a large sample size produces a better estimate of the population parameter. On the basis of sample size, the present meta-analysis was more likely to produce an estimate that was closer to the population parameter (Rosenthal, 1984). Furthermore, although Verbaten (2003) reported that significant heterogeneity existed in the aggregate effect size estimates produced in the analysis, few sensitivity analyses were conducted to identify possible sources of study-level selection bias.

The present meta-analysis conducted a number of sensitivity analyses to address the potential that inclusion criteria influenced the results. Finally, Verbaten (2003) questioned the possible causal relationship between ecstasy use and impaired cognitive performance on the basis that total lifetime cumulative ecstasy use was not found to be related to effect size in the previously conducted meta-analysis. No other possible dose-response relationships were investigated in the previous meta-analysis.

While the present meta-analysis also did not find a relationship between total lifetime cumulative ecstasy use and effect size, reported frequency of ecstasy use per month was identified as a statistically significant moderating variable. It is possible that the relationship between total lifetime cumulative ecstasy use and effect size was attenuated by low reliability in self-reported lifetime cumulative use. Although the reliability of self-reported total lifetime cumulative use was not tested in the present study, it is possible that participants were able to better estimate how much ecstasy they typically consume within a month than they were to correctly estimate the total number of tablets they have taken. Consider that mean estimates of lifetime use were in the hundreds and, in one study, was more than 1000 tablets (see Table 2). Only a small number of studies included participants with total lifetime cumulative ecstasy use over 500 tablets. It is possible that a relationship would be found if more studies recruited participants with very high cumulative use. It is notable, however, that large effect sizes were found in studies with total lifetime cumulative ecstasy use as low as 10 tablets.

While the present meta-analysis is an improvement over the previously conducted meta-analysis, there exist some limitations to conclusions based on the present study. Heterogeneity existed in the aggregate effect size estimates. Investigation of potential moderators that could account for the heterogeneity was limited by the amount of information reported in the primary

studies. A number of strategies were used to reduce the chance of committing a Type I error. In so doing, the effect size estimate might have become attenuated.   Conversely, as only published studies were used in the meta-analysis, the effect size might be overestimated. The effect size may also have been influenced by unmeasured exogenous variables. Finally, lack of random assignment to conditions may have created group selection bias. Discussion of the limitations follows.

Heterogeneity statistics revealed a significant amount of variability in the effect sizes from the included primary studies. Inspection of the Forrest plots for main and sub-analyses revealed that much of the variability was in the magnitude of the effect sizes, as opposed to the direction of the effect sizes. If the variance in the effect sizes was due to participant characteristics, then the true population effect will be different for different studies. If the variance was due to study methodology, then there may not be differences in true effect sizes. The differences may be due to bias (The Cochrane Collaboration, 2005). Sensitivity analyses were conducted to test if the effect would be different if more stringent exclusion criteria were used. The majority of the sensitivity analyses are described earlier in this paper.

Meta-regressions were also computed to see if the variance in effect sizes was associated with amount of ecstasy use or by duration of abstinence before testing. A limitation of testing the influence of possible moderator variables is that moderator variables in studies tend to be correlated and confounded with one another, making it difficult to disentangle meaningful relationships from spurious findings. The observational nature of meta-analysis, most problematic due to lack of random assignment by the meta-analyst of moderating variables to different primary studies, limits the certainty of conclusions (Lipsey, 2003).   Another difficulty is that moderating variables tend to not be adequately reported in primary studies, as is the case

40

in the many of the studies included in the present analysis (see Table 2). The more separate tests of the influence of potential moderators, the greater the chance of making a Type I error. This issue can sometimes be addressed by conducting a multivariate meta-regression (The Cochrane Collaboration, 2005). In the present study, multivariate moderator analysis was not tenable, due to uneven reporting of moderator information across the primary studies. This is a common difficulty in meta-analysis (Lipsey, 2003), because the meta-analyst does not have control over how the primary studies are conducted or reported. The increased chance of Type I error was averted by the use of a Bonferroni adjustment. Bonferroni adjustments can, however, increase the chance of a Type II error (Sankoh et al., 1997). Despite the increased chance of a Type II error, frequency of ecstasy use per month emerged as a moderating variable.

A number of other strategies were used to reduce the chance of Type I error in the meta-analysis. For example, the present procedure took the conservative approach of averaging the dependent outcomes (Rosenthal & DiMatteo, 2001). In interpretation, it should be remembered that the effect size is likely attenuated, however, the risk of Type I error is low. It may be useful in the future to use Rosenthal and Rubin's (1986) method of combining stochastically dependent outcomes and compare the results to the present analysis. The less conservative method of aggregation might result in a somewhat higher effect size estimate. However, without the original data, it would be difficult to discern which effect size estimate is a better parameter estimate.

In addition to decisions about how to aggregate the data from the included primary studies, search strategy decisions may have influenced the aggregate effect size estimate. Pooled effect sizes can differ depending on what studies are included or excluded from the aggregation. Sensitivity analyses were conducted to investigate if the effect size was excessively influenced

by the results of one included study. Results of sensitivity analyses suggest that the overall effect size for combined cognitive domains remained relatively stable when any one study was removed. Sub-analyses for individual cognitive domains were similarly stable when one study was removed, with a few exceptions (see Tables 10-19).

Despite an extensive search, it is possible that some published studies were not found. However, there is no reason to believe that the retrieved studies are not representative of the population of studies published in peer reviewed journals. Inspection of the funnel plot suggested that publication bias against non-significant results in small studies may have inflated the aggregate effect size. Therefore, Rosenthal's (1979) fail safe N was computed. The fail safe N was substantially large and suggests that readers can be confident that a relationship exists between ecstasy use and cognitive performance. For every study included in the meta-analysis, approximately 23 studies with an effect size equal to zero would need to exist in order to conclude that there was no statistically significant relationship between ecstasy use and cognitive performance.

There are some limitations of the fail safe N that should be noted, however. The fail safe N provides as estimate of the confidence that readers can have in the assertion that the effect size is not zero. It does not provide an estimate of the degree to which the effect size would remain clinically significant if all unpublished studies were included in the analysis. Another limitation is that the fail safe N assumes that the unpublished studies have an effect size equal to zero. It may be that some of the unpublished studies have an effect size that is in the opposite direction, with ecstasy users performing better on cognitive measures than ecstasy naïve controls. If that is the case, the fail safe N overestimates the number of studies that would be needed to bring the *p*-value above .05 (Begg, 1994). However, it should be noted that there may be a number of

unpublished, small studies in which the effect size was greater than zero, but not sufficiently greater to be statistically significant.   If that is the case, the fail safe N underestimates the number of studies that would be needed to bring the *p*-value above .05.   Readers should recognize that the fail safe N is an estimate.

The results of the present meta-analysis could be said to be a reasonable estimate of the relationship between recreational ecstasy use and cognitive performance as has been measured in the extant literature. Statistical covariation, however, is not itself a strong argument for a causal relationship between ecstasy use and cognitive performance. In the absence of randomized controlled trials, it is difficult to say the extent to which the results of the present meta-analysis approximate the true effect size for the neurocognitive consequences of ecstasy use. A meta-analysis comparing meta-analyses of randomized controlled trials versus meta-analyses of nonrandomized studies found little difference in the resultant overall effect size estimates (Lipsey & Wilson, 1993). Another meta-analysis, however, found that the degree to which the effect size estimates differ is dependent on the quality of the nonrandomized studies (Heinsman & Shadish, 1996).   A number of methodological quality-related variables were identified in the present meta-analysis. Due to the small number of primary studies that could be included in each cell, many methodological variables were not tested as potential moderators. However, sensitivity analyses suggested that these variables could affect the magnitude of the aggregate effect size, but did not appear to influence the direction. Even when studies used in meta-analysis are randomized controlled trials, meta-analysis is an observational study design, because the meta-analyst is not randomly assigning studies to different levels of independent variables. As with any observational study, the certainty of conclusions derived from meta-analysis is limited by the possibility of unmeasured, exogenous variables (Lipsey, 2003).

Ambiguous temporal precedence is another obvious threat to establishing a causal relationship using observational studies. It may be that premorbid impaired cognitive functioning exists in people who are more likely to use substances and that cognitive impairment hampers the decision-making process, leading people to select an immediately rewarding experience without consideration of possible negative consequences. It is also possible that there is a reciprocal causal relationship between cognitive processes and substance use. More than half of the included primary studies balanced the ecstasy use group and ecstasy naïve controls on general intelligence measures thought to be less likely to be impacted by lesions (see Table 4). Still, cognitive processes that have been associated with impulsivity and sensation-seeking have been found to be impacted by serotonergic dysfunction (Lezak, 1995) and trait impulsivity predicts escalation in drug use in rats (Robbins, 2005). Even though participants may be balanced on measures not usually impacted by lesions, their pre-morbid functioning within the cognitive domains hypothesized to be impaired by ecstasy use may be worse than that of controls. Prospective, longitudinal studies that measure serotonin-related cognitive processes before participants begin using ecstasy might help to disentangle the potential confounds.

A FDA-approved, Phase II, randomized double-blind placebo-controlled pilot study of MDMA is currently being conducted. The researchers are testing MDMA as a possible adjunct to psychotherapy in participants with posttraumatic stress disorder. Cognitive testing is being conducted before and after administration of MDMA (Mithoefer, 2003). The sample has limited generalizability and it is unlikely that the sample size for the clinical trial will be sufficient for the size of the effect of MDMA on cognitive performance to be statistically significant. Furthermore, the amount of MDMA provided to participants might not be large enough to produce a measurable effect, particularly considering the moderating relationship found in the

44

present meta-analysis between frequency of ecstasy use per month and effect size. Still, it is hoped that the researchers will conduct follow-up tests of the possible long-term cognitive sequelae of MDMA use.

The moderate magnitude of the effect size found in the present meta-analysis and the consistency of the direction of the effect across a variety of samples is enough to caution researchers conducting studies on the therapeutic value to pay special attention to the margin of safety from a cognitive standpoint. In addition, the present meta-analysis suggests that researchers conducting clinical trials of the safety of the use of MDMA should focus part of their assessment on measures of verbal memory. Measures of verbal memory produced the largest effect sizes that continued to be statistically significant during the one study removed sensitivity analyses. It is possible that measures of verbal memory would be the most sensitive to potential MDMA-related lesions. However, given the results and the limitations of the present meta-analysis, other measures of cognitive performance should also be included so as not to spuriously conclude that ecstasy does not cause long term cognitive impairment.

Finally, results of this meta-analysis may be used to help educate people who may be considering recreational ecstasy use. It is not recommended that the public be told that their cognitive performance will decrease by one-half a standard deviation if they consume ecstasy. It is, however, appropriate to report that there exists within the extant published empirical literature a moderate relationship between relatively impaired cognitive performance and recreational ecstasy use. The relationship is strongest for verbal memory, but small to moderate relationships have also been found in nonverbal memory, reaction time, fluency, attention, and executive functioning. Findings suggest that more frequent ecstasy use is associated with greater cognitive impairment, cognitive impairment can occur after relatively low amounts of total lifetime

45

cumulative use, and recovery of functioning does not appear to occur within one year post

cessation.

LIST OF REFERENCES

* indicates that the study was used in the meta-analysis

* Back-Madruga, C., Boone, K.B., Chang, & L. (2003). Neuropsychological Effects of 3,4-
Methylenedioxymethamphetamine (MDMA or Ecstasy) in Recreational Users. *Clinical Neuropsychologist, 17,* (4), 446-459.

Begg, C.B. (1994). Publication bias. In H. Cooper & L. V. Hedges (Eds.) *The handbook of research synthesis*. New York: Russell Sage Foundation.

Begg, C. B. & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50,* 1088–1101.

Barrowman, H.J., Fang, M., Sampson, M., & Moher, D. (2003). Identifying null meta-analyses that are ripe for updating. *BioMed Central Medical Research Methodology, 3,* 13-21.

* Bhattachary, S. & Powell, J.H. (2001). Recreational use of 3,4-
methylenedioxymethamphetamine (MDMA) of "ecstasy': Evidence for cognitive impairment. *Psychological Medicine, 31,* (4), 647-658.

* Bond, A.J., Verheyden, S.L., Wingrove, J., & Curran, H.V. (2004). Angry cognitive bias, trait aggression and impulsivity in substance users. *Psychopharmacology, 171,* (3), 331-339.

Boot, B.P., McGregor, I.S., & Hall, W. (2000). MDMA (ecstasy) neurotoxicity: Assessing and communicating risks. *Lancet, 355,* 1818-1825.

Buchanan, J.F. & Brown, C.R. (1988). 'Designer drugs'. A problem in clinical toxicology. *Medical Toxicology and Adverse Drug Experience, 3,* (1), 1-17.

Buhot, M. (1997). Serotonin receptors in cognitive behaviors. *Current Opinion Neurobiology, 7,* 243-254.

Chiarello, R. J. & Cole, J. O. (1987). The use of psychostimulants in general psychiatry: A

    reconsideration. *Archives of General Psychiatry, 44,* (3), 286-295.

Cohen. J. (1988). *Statistical power analysis for the behavioral sciences,2nd ed.* Hillsdale:

    Lawrence Erlbaum Associates.

* Croft, R.J., Klugman, A., Baldeweg, T., & Gruzelier, J.H. (2001). Electrophysiological

    evidence of serotonergic impairment in long-term MDMA ("ecstasy") users. *American*

    Journal of Psychiatry, 158, (10),1687-1692.

Curran, H. V. (2000). Is MDMA ('Ecstasy') Neurotoxic in Humans? An Overview of Evidence

    and of Methodological Problems in Research. *Neuropsychobiology, 42*, (1), 34-42.

* Curran, H.V. & Verheyden, S.L. (2003). Altered response to tryptophan supplementation

    after long-term abstention from MDMA (ecstasy) is highly correlated with human

    memory function. *Psychopharmacology, 169,* (1), 91-103.

Dafters, R. I., Hoshi, R., & Talbot, A.C. (2004). Contribution of cannabis and MDMA (ecstasy)

    to cognitive changes in long-term polydrug users. *Psychopharmacology, 173,* 405-410.

* Daumann, J., Fimm, B., & Willmes, K. (2003). Cerebral activation in abstinent ecstasy

    (MDMA) users during a working memory task: A functional magnetic resonance

    imaging (fMRI) study. *Cognitive Brain Research, 16,* (3), 479-487.

* Daumann, J., Fischermann, T., Pilatus, U., Thron, A,. Moeller-Hartmann, W., & Gouzoulis-

    Mayfrank, E. (2004). Proton magnetic resonance spectroscopy in ecstasy (MDMA)

    users. Neuroscience Letters, 362,113-116.

Fischer, C., Hatzidimitriou, G., Wlos, J., Katz, J., & Ricaurte, G. (1995). Reorganization of
ascending 5-HT axon projections in animals previously exposed to the recreational drug
+/- 3,4-methylenedioxymethamphetamine (MDMA, ecstasy). *Journal of Neuroscience,
15,* 5476-5485.

* Fisk, J.E., Montgomery, C, & Murphy, P. (2004). Evidence for executive deficits among users
of MDMA (ecstasy). *British Journal of Psychology, 95,* (4), 457-466.

* Fox, H. C., McLean, A., & Turner, J. J. D. (2002). Neuropsychological evidence of a
relatively selective profile of temporal dysfunction in drug-free MDMA ("ecstasy")
polydrug users. *Psychopharmacology, 162,* (2), 203-214.

* Fox, H.C., Parrott, A.C., & Turner, J.J.D. (2001). Ecstasy use: Cognitive deficits related to
dosage rather than self-reported problematic use of the drug. *Journal of
Psychopharmacology, 15,* (4), 273-281.

* Gouzoulis-Mayfrank, E, Daumann, J, Tuchtenhagen, F, Pelz, S, Becker, S, Kunert, H.J., Fimm,
B, & Sass, H. (2000). Impaired cognitive performance in drug-free recreational ecstasy
(MDMA) users. *Journal of Neurology, Neurosurgery, and Psychiatry, 68,* 719-725.

* Gouzoulis-Mayfrank, E., Thimm, B, & Rezk, M. (2003). Memory impairment suggests
hippocampal dysfunction in abstinent ecstasy users. *Progress in Neuro-
Psychopharmacology & Biological Psychiatry, 27,* (5), 819-827.

Gouzoulis-Mayfrank, E., Fischermann, T., Rezk, M., Thimm, B., Hensen, G., & Daumann, J.
(2005). Memory performance in polyvalent MDMA (ecstasy) users who continue or
discontinue MDMA use. *Drug and Alcohol Dependence, 78,* 317-323.

* Gouzoulis-Mayfrank, E., Thimm, B., & Rezk, M. (2003). Memory impairment suggests

hippocampal dysfunction in abstinent ecstasy users. *Progress in Neuro-*

Psychopharmacology & Biological Psychiatry, 27, (5), 819-827.

Grant, I., Gonzalez, R., Carey, C.L., Natarajan, L., & Wolfson, T. (2003). Non-acute (residual)

neurocognitive effects of cannabis use: A meta-analysis. *Journal of International*

Neuropsychological Society, 9, 679-689.

Green, A. R. (2004). MDMA: Fact and fallacy, and the need to increase knowledge in both the

scientific and popular press. *Psychopharmacology, 173,* 231-233.

Greenhouse, J. B. & Iyengar, S. (1994). Sensitivity analysis and diagnostics. In H. Cooper & L.

V. Hedges (Eds.) *The handbook of research synthesis*. New York: Russell Sage

Foundation.

Greer, G.R. & Tolbert, R. (1998). A method of conducting therapeutic sessions with MDMA.

Journal of Psychoactive Drugs, 30, (4), 371-379.

Grinspoon, L. & Bakalar, J.B. (1986). Can drugs be used to enhance the psychotherapeutic

process? *American Journal of Psychotherapy, 40,* (3), 393-404.

* Halpern, J.H., Pope, H.G., Sherwood, A.R., Barry, S., Hudson, J.I., & Yurgelun-Todd, D.

(2004). Residual neuropsychological effects of illicit 3,4-

methylenedioxymethamphetamine (MDMA) in individuals with minimal exposure to

other drugs. Drug & Alcohol Dependence, 75, (2),135-147.

* Hanson, K.L. & Luciana, M. (2004). Neurocognitive function in users of MDMA: The

importance of clinically significant patterns of use. *Psychological Medicine, 34,* (2), 229-

246.

Hatzidimitriou, G., McCann, U.D., & Ricaurte, G. A. (1999). Altered serotonin innervation

    patterns in the forebrain of monkeys treated with ($\pm$)3,4,-

    methylenedioxymethamphetamine seven years previously: Factors influencing abnormal

    recovery. *Journal of Neuroscience,19,* (12), 5096-5017.

* Heffernan, T.M., Jarvis, H., & Rodgers, J. (2001). Prospective memory, everyday cognitive

    failure and central executive function in recreational users of Ecstasy. *Human*

    Psychopharmacology: Clinical & Experimental, 16, (8), 607-612.

Heinsman, D.T. & Shadish, W.R. (1996). Assignment methods in experimentation: When do

    nonrandomized experiments approximate answers from randomized experiments?

    Psychological Methods, 1, (2), 154-169.

Hedges, L.V. (1981). Distribution theory for Glass's estimator of effect size and related

    estimators. *Journal of Educational Statistics, 6,* 107-128.

Hedges, L.V. (1994). Fixed effects models. In H. Cooper & L. V. Hedges (Eds.) *The handbook*

    of research synthesis. New York: Russell Sage Foundation.

Hedges, L.V. & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic.

Higgins, J.P.T. & Thompson, S. G. (2002). Quantifying heterogeneity in meta-analysis.

    Statistics in medicine, 21, 1539-1558.

Hunter, J.E., Schmidt, F.L., & Jackson, G.B. (1982). *Meta-analysis: Cumulating research*

    findings across studies. Beverly Hills: Sage.

Hurley, R.A., Reneman, L., & Taber, K.H. (2002). Ecstasy in the brain: A model for

    neuroimaging. *Journal of Neuropsychiatry and Clinical Neuroscience, 14,* (2), 125-129.

Imperio, W.A. (2001). Ecstasy research okayed. *Clinical Psychiatry News, 29,* 44.

* Jacobsen, L.K., Mencl, W.E., Kenneth, P.R., Skudlarkshi, P., & Krystal, J.H. (2004). Preliminary evidence of hippocampal dysfunction in adolescent MDMA ("ecstasy") users: possible relationship to neurotoxic effects. *Psychopharmacology, 173*, (3-4), 383-390.

Kish, S. J. (2002). How strong is the evidence that brain serotonin neurons are damaged in human users of ecstasy? *Pharmacology, Biochemistry & Behavior, 71,* (4), 845-855.

Kulik, J. (1983). Review of G.V. Glass et al. (1981). *Evaluation News, 4,* 101-105.

Lezak, M. D. (1995). *Neuropsychological assessment (3rd Ed.).* New York: Oxford Press.

Lipsey, M.W. (2003). Those confounding moderators in meta-analysis: Good, bad, and ugly. *The annals of the American academy, 587,* 69-81.

Lipsey, M.W. & Wilson, D.B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48,* 1181-1209.

Lipsey, M.W. & Wilson, D.B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Luciana, M., Burgund, E.D., Berman, M., & Hanson, K.L. (2001). Effects of tryptophan loading on verbal, spatial, and affective working memory functions in healthy adults. *Journal of Psychopharmacology, 15,* 219-230.

Lyles, J. & Cadet, J.L. (2003). Methylenedioxymethamphetamine (MDMA, ecstasy) neurotoxicity: Cellular and molecular mechanisms. *Brain Research Reviews, 42,* 155-168.

Marin-Martinez, F. & Sanchez-Meca, J. (1999). Averaging dependent effect sizes in meta-analysis: a cautionary note about procedures. *Spanish Journal of Psychology, 2,* (1), 32-8.

* McCann, U.D., Mertl, M., Eligulashvili, V., & Ricaurte, G.A. (1999). Cognitive performance in +/- 3, 4 methylenedioxymethamphetamine (MDMA, ecstasy) users: A controlled study. *Psychopharmacology, 143,* 417-425.

McCann, U.D., Ridenour, A., Shaham, Y., & Ricaurte, G.A. (1994). Brain serotonergic neurotoxicity after MDMA (ecstasy): A controlled study in humans. *Neuropsychopharmacology, 10*, 129-138.

McCann, U.D., Szabo, Z., Scheffel, U., Dannals, R.F., & Ricaurte, G.A. (1998). Positron emission tomographic evidence of toxic effect of MDMA ("Ecstasy") on brain serotonin neurons in human beings. *Lancet, 352,* 1433-1437.

* McCardle, K., Luebbers, S., & Carter, J. D. (2004). Chronic MDMA (ecstasy) use, cognition and mood. *Psychopharmacology, 173*, (3), 434-439.

Mithoefer, M. C. (2003). *Study protocol: Phase II clinical trial testing the safety and efficacy of 3, 4-methylenedioxymethamphetamine (MDMA)-assisted psychotherapy in subjects with chronic posttraumatic stress disorder*, Study # 63-384.

* Montgomery, C., Fisk, J.E., Newcombe, R., Wareing, M., & Murphy, P.N. (2005). Syllogistic reasoning performance in MDMA (ecstasy) users. *Experimental & Clinical Psychopharmacology, 13,* (2), 137-145.

Montoya, A. G., Sorrentino, R., Lukas, S. E., & Price, B.H. (2002). Long-term neuropsychiatric consequences of "Ecstasy" (MDMA): A review. *Harvard Review of Psychiatry, 10,* (4), 212-220.

* Morgan, M.J. (1998). Recreational use of "ecstasy" (MDMA) is associated with elevated impulsivity. *Neuropsychopharmacology, 19,* (4), 252-264.

* Morgan, M. J. (1999). Memory deficits associated with recreational use of "ecstasy" (MDMA). *Psychopharmacology, 141*, (1), 30-36.

Morgan, M. J. (2000). Ecstasy (MDMA): A review of its possible persistent psychological effects. *Psychopharmacology, 152,* (3), 230-248.

Office of Applied Studies, Substance Abuse and Mental Health Services Administration. (2005). *Reports from the National Survey on Drug Use and Health*. Substance Abuse and Mental Health Services Administration. Online at http://oas.samhsa.gov/nsduh.htm#Reports.

Ozer, E. J., Best, S. R., Lipsey, T. L., & Weiss, D. S. (2003). Predictors of posttraumatic stress disorder and symptoms in adults: A meta-analysis. *Psychological Bulletin, 129,* (1), 52-73.

* Parrott, A.C., Lees, A., & Garnham, N. J. (1998). Cognitive performance in recreational users of MDMA or 'ecstasy': Evidence for memory deficits. *Journal of Psychopharmacology, 12,* (1), 79-83.

Parrott, A.C. (2000). Human research on MDMA (3,4-methylene- dioxymethamphetamine) neurotoxicity: Cognitive and behavioural indices of change. *Neuropsychobiology, 42,* (1), 17-25.

Parrott, A.C. (2002). Recreational ecstasy/MDMA, the serotonin syndrome, and serotonergic neurotoxicity. *Pharmacology, Biochemistry, & Behavior, 71,* (4), 837-44.

Parrott, A. C., Buchanan, T., Scholey, A. B., Heffernan, T., Ling, J., & Rodgers, J. (2002). Ecstasy/MDMA attributed problems reported by novice, moderate and heavy recreational users. *Human Psychopharmacology: Clinical & Experimental, 17,* (6), 309-312.

Parrott, A. C., Gouzoulis-Meyfrank, E., & Rodgers, J. (2004). Ecstasy/MDMA and cannabis: The complexities of their interactive neuropsychobiological effects. *Journal of Psychopharmacology, 18,* (4), 572-575.

Peroutka, S.J., Pascoe, N., & Faull, K.F. (1987). Monoamine metabolites in the cerebrospinal fluid of recreational users of 3,4-methylenedioxymethamphetamine (MDMA, ecstasy). *Research Communication in Substances of Abuse, 8,* 125-138.

Pope, H.G., Gruber, A.J., Hudson, J.I., Huestis, M.A., & Yurgelun-Todd, D. (2001). Neuropsychological performance in long-term cannabis users. *Archives of General Psychiatry, 58,* 909-915.

Rabbitt, P. (1997). *Methodology of frontal and executive functions.* Psychology Press: Hove, E. Sussex.

Raudenbush, S. W. (1994). Random effects models. In H. Cooper & L. V. Hedges (Eds.) *The handbook of research synthesis*. New York: Russell Sage Foundation.

Reneman, L., Booij, J., Majoie, C., van den Brink, W., & den Heeten, G. J. (2001). Investigating the potential neurotoxicity of Ecstasy (MDMA): an imaging approach. *Human Psychopharmacology, 16,* 579-588.

* Reneman, L., Booij, J., & Schmand, B. (2000). Memory disturbances in "ecstasy" users are correlated with an altered brain serotonin neurotransmission. Psychopharmacology, 148, (3), 322-324.

* Reneman, L., Lavalaye, J., & Schmand, B. (2001). Cortical serotonin transporter density and verbal memory in individuals who stopped using 3.4-methylenedioxymethamphetamine (MDMA or "ecstasy"). *Archives of General Psychiatry, 58,* (10), 901-906.

* Reneman, L., Majoie, C.B.L.M., Schmand, B., van den Brink, W., & den Heeten, G. J.
(2001). Prefrontal N-acetylaspartate is strongly associated with memory performance in
(abstinent) ecstasy users: Preliminary report. *Society of Biological Psychiatry, 50*, (7),
550-554.

Ricaurte, G.A., DeLanney, L.E., Irwin, I., & Langston, J.W. (1988). Toxic effects of MDMA
on central serotonergic neurons in the primate: Importance of route and frequency of drug
administration. *Brain Research, 446,* 165-168.

Ricaurte, G.A., Finnegan, K.T., Irwin, I., & Langston, J.W. (1990). Aminergic metabolites in
cerebrospinal fluid of humans previously exposed to MDMA: Preliminary observations.
Annals of the New York Academy of Sciences, 600, 699-710.

Robbins, T. (2005). *Cognitive sequelae of drug abuse: Studies in humans and experimental
animals.* President's lecture at the College on Problems of Drug Dependence, Orlando,
FL.

* Rodgers, J. (2000). Cognitive performance amongst recreational users of "ecstasy".
Psychopharmacology, 151, (1), 19-24.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological
Bulletin, 86,* 638-641.

Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills: Sage
Publications.

Rosenthal, R. (1994). Parametric measures of effect size. In H. M. Cooper & L.V. Hedges
(Eds.) , *The Handbook of Research Synthesis* (pp. 231-244). New York: Russell Sage
Foundation.

Rosenthal, R. & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52,* 59-82.

Rosenthal, R. & Rubin, D.B. (1986). Meta-analytic procedures for combining studies with multiple effect sizes. *Psychological Bulletin, 99,* 400-406.

Sankoh, A.J., Huque, M.F., & Dubin, N. (1997). Some comments on frequently used multiple endpoint adjustment methods in clinical trials. *Statistics in medicine, 16,* 2529-2542.

Shadish W. R. & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.) *The handbook of research synthesis.* New York: Russell Sage Foundation.

Shulgin, A. T. (1986). The background and chemistry of MDMA. Journal of *Psychoactive Drugs, 18,* (4), 291-304.

* Simon, N.G., & Mattick, R.P. (2002). The impact of regular ecstasy use on memory function. *Addiction, 97,* (12), 1523-1529.

Solowij, N., Hall, W., & Lee, N. (1992). Recreational MDMA use in Sydney: A profile of "ecstasy" users and their experiences with the drug. *British Journal of Addiction, 87,* 1161-1172.

Steele, T. D., McCann, U. D., & Ricaurte, G. A. (1994). 3,4-Methylenedioxy-methamphetamine (MDMA, ecstasy): Phamacology and toxicology in animals and humans. *Addiction, 89,* 539-551.

The Cochrane Collaboration. (2005). *The Cochrane manual, issue 2.* Updated February 23, 2005. http://www.cochrane.org/admin/manual.htm

Thomasius, R., Petersen, K., Buchert, R., Andresen, B., Zapletalova, P., Wartberg, L., Nebeling, B., Schmoldt, A. (2003). Mood, cognition, and serotonin transporter availability in current and former ecstasy (MDMA) users. *Psychopharmacology, 167,* 85-96.

Verbaten, M. N. (2003). Specific memory deficits in ecstasy users? The results of a meta-analysis. *Human Psychopharmacology, 18*, (4), 281-90.

Verdejo-Garcia, A., Lopez-Torrecillas, F., Gimenez, C. O., Perez-Garcia, M. (2004). Clinical implications and methodological challenges in the study of the neuropsychological correlates of cannabis, stimulant, and opioid abuse. *Neuropsychology review, 14,* (1), 1-41.

* Verkes, R.J., Gijsman, H. J., Pieters, M.S.M., Schoemaker, R.C., de Visser, S., Kuijpers, M., Pennings, E.J.M., de Bruin, D., Van de Wijngaart, G., Van Gerven, Joop M.A., & Cohen, A. F. (2001). Cognitive performance and serotonergic function in users of ecstasy. *Psychopharmacology, 153,* (2), 196-103.

Vollenweider, F.X., Gamma, A., Liechti, M., & Huber, T. (1999). Is a single dose of MDMA harmless? *Neuropsychopharmacology, 21,* (4), 598-600.

* von Geusau, N.A., Stalenhoef, P., Huizinga, M., Snel, J., & Ridderinkhof, K.R. (2004). Impaired executive function in male MDMA ("ecstasy") users. *Psychopharmacology, 175,* (3), 331-341.

* Wareing, M., Fisk, J.E., & Murphy, P.N. (2000). Working memory deficits in current and previous users of MDMA (ecstasy). *British Journal of Psychology, 91*, (2), 181-188.

* Wareing, M., Fisk, J.E., & Murphy, P.N. (2005). Visuo-spatial working memory deficits in current and former users of MDMA ('ecstasy'). *Human Psychopharmacology: Clinical & Experimental, 20*, 115-123.

* Wareing, M., Murphy, P.N., & Fisk, J.E. (2004) Visuospatial memory impairments in users of

    MDMA ('ecstasy'). *Psychopharmacology, 173,* (3), 391-397.

Wolf, F.M. (1986). *Meta analysis: Quantitative methods for research synthesis*.

Beverly Hills, CA: Sage

* Yip, & Lee, (2005). Effect of ecstasy use on neuropsychological function: A study in Hong

    Kong. *Psychopharmacology, 179,* 620-628.

* Zakzanis, K.K., Young, D.A., & Radkhoshnoud, N.F. (2002). Attentional processes in

    abstinent methylenedioxymethamphetamine (Ecstasy) users. *Applied Neuropsychology,*

    9, (2), 84-91.

* Zakzanis, K.K. & Young, D.A. (2001). Memory impairment in abstinent MDMA (ecstasy)

    users: A longitudinal investigation. *Neurology, 56,* 966-969.

APPENDIX A
TABLES

Table 1: Sample size for each primary study

Study	Ecstasy n	Control n
Back-Madruga, et al., 2003	22	28
Bhattachary & Powell, 2001	16	20
Bond, et al., 2004	32	32
Croft, et al., 2001	11	31
Curran & Verheyden, 2003	32	32
Daumann, et al., 2003	11	11
Daumann, et al., 2004	13	13
Fisk, et al., 2004	44	59
Fox, et al., 2002	20	20
Fox, Parrott, et al., 2001	11	20
Gouzoulis-Mayfrank, et al., 2000	28	28
Gouzoulis-Mayfrank, et al., 2003	30	30
Halpern, et al., 2004	23	16
Hanson & Luciana, 2004	26	26
Heffernan, et al., 2001 study 2	30	37
Jacobsen, et al., 2004	6	6
McCann, et al., 1999	22	23
McCardle, et al., 2004	17	15
Montgomery, et al., 2005	22	26
Morgan, 1998 study 1	16	12
Morgan, 1998 study 2, 1999	25	20
Parrott, et al., 1998	10	10
Reneman, et al., 2000	5	9
Reneman, Majoie, et al., 2001	8	7
Reneman, Lavalaye, et al., 2001	16	13
Rodgers, 2000	15	15
Simon & Mattick, 2002	40	37
Verkes, et al., 2001	21	20
von Geusau, et al., 2004 men	17	12
von Geusau, et al., 2004 women	9	21
Wareing, et al., 2000	10	10
Wareing, et al., 2005	12	31
Wareing, Fisk, et al., 2004	17	31
Wareing, Murphy, et al., 2004	10	18
Yip & Lee, 2005	100	100
Zakzanis, et al., 2002	24	30
Zakzanis, et al., 2003	15	17
Total	786	886

Table 2: Summary of ecstasy use for groups of ecstasy users included in the meta-analysis from each study.

Study	Amount of days of required abstinence from ecstasy before testing	Mean frequency of ecstasy use per month	Mean cumulative lifetime number of ecstasy tablets consumed	Mean number of years of ecstasy use
Back-Madruga, et al., 2003	1		74.6	6.3
Bhattachary & Powell, 2001	30			
Bond, et al., 2004	365	7	1105.85	3.5
Croft, et al., 2001	7		41.9	
Curran & Verheyden, 2003	365	6.95		3.49
Daumann, et al., 2003	7	2.78	258.18	4.43
Daumann, et al., 2004	7	2.31	324.54	2.69
Fisk, et al., 2004	7	1.76	343.38	3.52
Fox, et al., 2002	14		172	4.33
Fox, Parrott, et al., 2001	14		918.2	5.38
Gouzoulis-Mayfrank, et al., 2000	7	3.5	93.4	2.25
Gouzoulis-Mayfrank, et al., 2003	7	4.5	503.2	3.41
Halpern, et al., 2004	10			
Hanson & Luciana, 2004	7	2.3	64.9	2.25
Heffernan, et al., 2001 study 2	1	5.6		
Jacobsen, et al., 2004			10	
McCann, et al., 1999	21	5.72	215	
McCardle, et al., 2004	7			2.2
Montgomery, et al., 2005	7	1.88	303.3	3.17
Morgan, 1998 study 1	3	2.94	35.6	2.12
Morgan, 1998 study 2, 1999	3	4.36	49.6	4.12
Parrott, et al., 1998				
Reneman, et al., 2000	60		218	
Reneman, Majoie, et al., 2001	7		902	6.6
Reneman, Lavalaye, et al., 2001	365		268	4.6
Rodgers, 2000	60		20	
Simon & Mattick, 2002	1	2.4	258	3.83
Verkes, et al., 2001	7		741	4.5
von Geusau, et al., 2004 men	14	1.96	53.82	2.28
von Geusau, et al., 2004 women	14	1.44	38.78	2.24
Wareing, et al., 2000	180	8.05		3.9
Wareing, et al., 2005	180		433	
Wareing, Fisk, et al., 2004	180		385	3.79
Wareing, Murphy, et al., 2004	180		469.2	4.4
Yip & Lee, 2005	1	14.8	35.84	0.2
Zakzanis, et al., 2002	14	1.7	22.3	14.7
Zakzanis, et al., 2003	14	2.4	19	1.53

Blank cells indicate that there was no information in reference to moderator included in the study manuscript.

Table 3: Other substance use characteristics of ecstasy use group and control group.

Study	Description of control group	Balanced for marijuana	Balanced for alcohol	Balanced for other drugs
Back-Madruga, et al., 2003				
Bhattachary & Powell, 2001	non-drug users	E>C	y	y
Bond, et al., 2004	mixed	y	y	
Croft, et al., 2001	non-drug users			
Curran & Verheyden, 2003	marijuana	y	y	E>C
Daumann, et al., 2003	non-drug users	E>C		E>C
Daumann, et al., 2004	marijuana	E>C		E>C
Fisk, et al., 2004	mixed	E>C	E>C	
Fox, et al., 2002	polydrug users	y	y	E>C
Fox, Parrott, et al., 2001	mixed	E>C	y	E>C
Gouzoulis-Mayfrank, et al., 2000	marijuana	y		
Gouzoulis-Mayfrank, et al., 2003	marijuana	E>C		E>C
Halpern, et al., 2004	mixed	E>C	E>C	
Hanson & Luciana, 2004				
Heffernan, et al., 2001 study 2	mixed			
Jacobsen, et al., 2004	mixed	y	y	y
McCann, et al., 1999	mixed	y		E>C
McCardle, et al., 2004	mixed	E>C	E>C	E>C
Montgomery, et al., 2005	mixed	E>C	E>C	E>C
Morgan, 1998 study 1	polydrug users	y	y	y
Morgan, 1998 study 2, 1999	polydrug users	y	y	y
Parrott, et al., 1998				
Reneman, et al., 2000	non-drug users			
Reneman, Majoie, et al., 2001	mixed			y
Reneman, Lavalaye, et al., 2001	mixed	E>C	y	E>C
Rodgers, 2000	marijuana			
Simon & Mattick, 2002	marijuana	y	y	E>C
Verkes, et al., 2001	mixed	E>C	y	E>C
von Geusau, et al., 2004 men	mixed	E>C		E>C
von Geusau, et al., 2004 women	mixed	E>C		E>C
Wareing, et al., 2000	non-drug users	E>C		E>C
Wareing, et al., 2005	mixed	E>C	C>E	E>C
Wareing, Fisk, et al., 2004	mixed			
Wareing, Murphy, et al., 2004	mixed			
Yip & Lee, 2005				
Zakzanis, et al., 2002	mixed			
Zakzanis, et al., 2003	polydrug users			

Blank cells in the description of the control group indicate that there was no information in the study manuscript in reference to the drug use of the control group. "Mixed" indicates that the control group contained non-drug users and drug users. Blank cells in the remainder of the columns indicate that the difference in use was not statistically tested in the study. A "y" in the cell indicates that no statistical significance was found between groups. "E>C" indicates that the ecstasy use group used significantly more than the control group. "C>E" indicates that the control group used significantly more than the ecstasy use group.

Table 4: Balance of demographic variables between ecstasy use group and control group.

Study	Balanced for IQ	Balanced for education	Balanced for age
Back-Madruga, et al., 2003	y	y	y
Bhattachary & Powell, 2001	y		y
Bond, et al., 2004	y		E>C
Croft, et al., 2001			
Curran & Verheyden, 2003	y		E>C
Daumann, et al., 2003		y	y
Daumann, et al., 2004		y	y
Fisk, et al., 2004	y	y	y
Fox, et al., 2002	y		y
Fox, Parrott, et al., 2001	y	C>E	y
Gouzoulis-Mayfrank, et al., 2000	C>E	y	y
Gouzoulis-Mayfrank, et al., 2003	C>E	y	y
Halpern, et al., 2004	y		C>E
Hanson & Luciana, 2004	y	y	y
Heffernan, et al., 2001 study 2			y
Jacobsen, et al., 2004	y	y	y
McCann, et al., 1999	y	C>E	C>E
McCardle, et al., 2004	y	y	y
Montgomery, et al., 2005	y	y	y
Morgan, 1998 study 1	y	y	y
Morgan, 1998 study 2, 1999	y	y	y
Parrott, et al., 1998			
Reneman, et al., 2000		y	y
Reneman, Majoie, et al., 2001	y	C>E	y
Reneman, Lavalaye, et al., 2001	y	C>E	y
Rodgers, 2000			
Simon & Mattick, 2002	C>E	y	y
Verkes, et al., 2001		y	y
von Geusau, et al., 2004 men		y	y
von Geusau, et al., 2004 women		y	y
Wareing, et al., 2000		y	y
Wareing, et al., 2005	y	C>E	y
Wareing, Fisk, et al., 2004	y	y	y
Wareing, Murphy, et al., 2004			
Yip & Lee, 2005	y	y	y
Zakzanis, et al., 2002	y	y	y
Zakzanis, et al., 2003	C>E	y	y

Blank cells indicate that the difference in the demographic variable was not statistically tested in the study. A "y" in the cell indicates that no statistical significance was found between groups. "E>C" indicates that the ecstasy use group mean was significantly higher than the control group on that variable. "C>E" indicates that the control group mean was significantly higher than the ecstasy use group on that variable.

Table 5: Combined estimated effect size using Hedge's *g*

		Effect size and 95% confidence interval					Test of null (2-Tail)	
Model	Number Studies	Point estimate	Standard error	Variance	Lower limit	Upper limit	*Z*-value	*P*-value
Fixed effects	37	-0.498	0.052	0.003	-0.600	-0.395	-9.547	0.0000
Random effects	37	-0.515	0.083	0.007	-0.678	-0.353	-6.220	0.0000

Table 6: Heterogeneity statistics

Estimate	Q-value	df (Q)	P-value	I^2
Combined cognitive domains	84.252	36	0.000	57.271
Cognitive domains analyzed separately				
Attention	61.932	21	0.000	66.092
Immediate verbal memory	80.376	25	0.000	68.896
Delayed verbal memory	273.306	17	0.000	93.780
Immediate nonverbal memory	35.888	19	0.011	47.058
Delayed nonverbal memory	77.765	9	0.000	88.427
Executive functioning	15.352	12	0.223	21.835
Fluency	22.226	9	0.008	59.507
Reaction time	25.845	11	0.004	61.307
Learning	4.55	7	0.714	0.000
Control group consists of				
marijuana users	1.890	5	0.864	0.000
non-drug users	7.013	4	0.135	42.962
polydrug users	0.065	3	0.996	0.000
mixed	39.813	17	0.001	57.300
Including only studies with control groups balanced for alcohol, marijuana, and other substance use	0.957	3	0.620	0.000
Including only studies with control groups balanced for intelligence, education, and age	38.443	11	0.000	73.987

Table 7: Aggregate effect sizes for sub-analyses by cognitive domain

		Number of studies	Hedge's g	Standard error	Z-value	P-value
Attention	Random	22	-0.297	0.113	-2.625	0.009
	Fixed		-0.296	0.063	-4.726	0.000
Immediate verbal memory	Random	26	-0.634	0.112	-5.653	0.000
	Fixed		-0.640	0.061	-10.542	0.000
Delayed verbal memory	Random	18	-0.880	0.317	-2.773	0.006
	Fixed		-0.830	0.078	-10.621	0.000
Immediate nonverbal memory	Random	20	-0.428	0.095	-4.511	0.000
	Fixed		-0.427	0.066	-6.501	0.000
Delayed nonverbal memory	Random	10	-0.634	0.265	-2.390	0.017
	Fixed		-0.667	0.086	-7.729	0.000
Executive functioning	Random	13	-0.238	0.100	-2.385	0.017
	Fixed		-0.226	0.086	-2.626	0.009
Fluency	Random	10	-0.452	0.140	-3.220	0.001
	Fixed		-0.455	0.085	-5.342	0.000
Reaction time	Random	11	-0.492	0.178	-2.756	0.006
	Fixed		-0.431	0.107	-4.027	0.000
Learning	Random	8	-0.151	0.112	-1.349	0.177
	Fixed		-0.151	0.112	-1.349	0.177

The Bonferroni adjusted one-tailed $Z_{critical}$ = -2.128. The Bonferroni adjustment was modified for correlated outcomes.

Table 8: Mixed effects regression (unrestricted maximum likelihood) of potential moderator variables on effect size.

Variable	Number Studies	Slope point estimate	Q model	Q residual	Q total	Z-value	P-value
Total lifetime cumulative ecstasy use	30		0.17	30.21	30.39	-0.417	0.6768
Frequency of ecstasy use per month	20	-0.091	33.60	27.44	61.04	-5.797	0.0000
Number of years used ecstasy	27		3.22	25.73	28.95	1.79	0.0729
Number of years used ecstasy, with one extreme score removed	26		2.11	25.86	27.98	1.45	0.1460
Duration of required abstinence from ecstasy before cognitive testing	35		0.33	35.26	35.59	-0.572	0.5670

Slope point estimates are included only for variables that are related to a significant amount of variance in effect sizes. Bonferroni adjusted $Z_{critical}$ = +/- 2.5688.

Table 9: Sensitivity analysis. Aggregate effect sizes for sub-analyses of type of control group using random effects model

		Number of studies	Hedge's *g*	Standard error	*Z*-value	*P*-value
Control group consists of	non-drug users	5	-0.831	0.254	-3.270	0.001
	mixed	18	-0.571	0.122	-4.679	0.000
	marijuana users	6	-0.345	0.118	-2.924	0.003
	polydrug users	4	-0.250	0.165	-1.511	0.131
Including only studies with control groups balanced for alcohol, marijuana, and other substance use		3	-0.338	0.217	-1.558	0.119
Including only studies with control groups balanced for intelligence, education, and age		11	-0.348	0.169	-2.061	0.039

Table 10: Combined estimated effect size using Hedge's *g* and the random effects model with one study removed.

	Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-Value	*p*-Value
Back-Madruga, et al., 2003	-0.536	0.083	0.007	-0.698	-0.374	-6.468	0.000
Bhattachary & Powell, 2001	-0.494	0.082	0.007	-0.655	-0.332	-5.997	0.000
Bond, et al., 2004	-0.527	0.085	0.007	-0.694	-0.361	-6.196	0.000
Croft, et al., 2001	-0.515	0.085	0.007	-0.682	-0.349	-6.063	0.000
Curran & Verheyden, 2003	-0.521	0.085	0.007	-0.687	-0.354	-6.120	0.000
Daumann, et al., 2003	-0.525	0.084	0.007	-0.690	-0.360	-6.227	0.000
Daumann, et al., 2004	-0.508	0.084	0.007	-0.674	-0.343	-6.021	0.000
Fisk, et al., 2004	-0.533	0.084	0.007	-0.698	-0.369	-6.348	0.000
Fox, et al., 2002	-0.522	0.085	0.007	-0.689	-0.356	-6.141	0.000
Fox, Parrott, et al., 2001	-0.521	0.085	0.007	-0.687	-0.355	-6.147	0.000
Gouzoulis-Mayfrank, et al., 2000	-0.522	0.086	0.007	-0.689	-0.354	-6.100	0.000
Gouzoulis-Mayfrank, et al., 2003	-0.522	0.086	0.007	-0.690	-0.354	-6.098	0.000
Halpern, et al., 2004	-0.519	0.085	0.007	-0.686	-0.352	-6.091	0.000
Hanson & Luciana, 2004	-0.529	0.085	0.007	-0.695	-0.363	-6.245	0.000
Heffernan, et al., 2001 study 2	-0.503	0.085	0.007	-0.669	-0.337	-5.938	0.000
Jacobsen, et al., 2004	-0.510	0.084	0.007	-0.675	-0.346	-6.080	0.000
McCann, et al., 1999	-0.481	0.079	0.006	-0.636	-0.326	-6.101	0.000
McCardle, et al., 2004	-0.521	0.085	0.007	-0.687	-0.354	-6.131	0.000
Montgomery, et al., 2005	-0.519	0.085	0.007	-0.686	-0.351	-6.072	0.000
Morgan, 1998 study 1	-0.522	0.085	0.007	-0.688	-0.357	-6.173	0.000
Morgan, 1998 study 2, 1999	-0.524	0.085	0.007	-0.691	-0.357	-6.163	0.000
Parrott, et al., 1998	-0.520	0.084	0.007	-0.686	-0.355	-6.158	0.000
Reneman, et al., 2000	-0.500	0.082	0.007	-0.661	-0.338	-6.062	0.000
Reneman, Majoie, et al., 2001	-0.507	0.084	0.007	-0.672	-0.343	-6.050	0.000
Reneman, Lavalaye, et al., 2001	-0.500	0.083	0.007	-0.664	-0.337	-5.992	0.000
Rodgers, 2000	-0.522	0.085	0.007	-0.688	-0.356	-6.155	0.000
Simon & Mattick, 2002	-0.528	0.085	0.007	-0.695	-0.361	-6.186	0.000
Verkes, et al., 2001	-0.505	0.085	0.007	-0.671	-0.340	-5.977	0.000
von Geusau, et al., 2004 men	-0.503	0.083	0.007	-0.667	-0.340	-6.040	0.000
von Geusau, et al., 2004 women	-0.536	0.082	0.007	-0.696	-0.376	-6.556	0.000
Wareing, et al., 2000	-0.508	0.084	0.007	-0.673	-0.343	-6.041	0.000
Wareing, et al., 2005	-0.501	0.084	0.007	-0.666	-0.337	-5.979	0.000
Wareing, Fisk, et al., 2004	-0.519	0.085	0.007	-0.686	-0.351	-6.078	0.000
Wareing, Murphy, et al., 2004	-0.508	0.084	0.007	-0.673	-0.343	-6.020	0.000
Yip & Lee, 2005	-0.466	0.074	0.005	-0.611	-0.321	-6.314	0.000
Zakzanis, et al., 2002	-0.534	0.084	0.007	-0.698	-0.370	-6.391	0.000
Zakzanis, et al., 2003	-0.525	0.085	0.007	-0.691	-0.359	-6.204	0.000
Overall ES	-0.515	0.083	0.007	-0.678	-0.353	-6.220	0.000

Information listed on row of named study represents the aggregate estimate with that study removed.

Table 11: Aggregate estimated effect size for measures of attention using Hedge's *g* and the random effects model with one study removed.

Study	ES	Standard error	Variance	Lower limit	Upper limit	Z-Value	*p*-Value
Back-Madruga, et al., 2003	-0.318	0.117	0.014	-0.548	-0.089	-2.724	0.006
Bond, et al., 2004	-0.303	0.120	0.014	-0.538	-0.069	-2.533	0.011
Croft, et al., 2001	-0.277	0.116	0.014	-0.505	-0.049	-2.384	0.017
Curran & Verheyden, 2003	-0.299	0.119	0.014	-0.532	-0.066	-2.514	0.012
Fox, et al., 2002	-0.297	0.118	0.014	-0.529	-0.065	-2.508	0.012
Gouzoulis-Mayfrank, et al., 2000	-0.300	0.119	0.014	-0.534	-0.066	-2.515	0.012
Gouzoulis-Mayfrank, et al., 2003	-0.308	0.119	0.014	-0.541	-0.074	-2.583	0.010
Halpern, et al., 2004	-0.288	0.118	0.014	-0.519	-0.057	-2.441	0.015
Hanson & Luciana, 2004	-0.302	0.119	0.014	-0.535	-0.069	-2.537	0.011
Jacobsen, et al., 2004	-0.295	0.116	0.013	-0.522	-0.067	-2.541	0.011
McCann, et al., 1999	-0.223	0.072	0.005	-0.363	-0.083	-3.113	0.002
McCardle, et al., 2004	-0.294	0.118	0.014	-0.524	-0.063	-2.495	0.013
Morgan, 1998 study 1	-0.290	0.117	0.014	-0.520	-0.061	-2.477	0.013
Morgan, 1998 study 2, 1999	-0.297	0.119	0.014	-0.529	-0.064	-2.504	0.012
Parrott, et al., 1998	-0.305	0.117	0.014	-0.534	-0.077	-2.619	0.009
Rodgers, 2000	-0.306	0.117	0.014	-0.536	-0.076	-2.607	0.009
Simon & Mattick, 2002	-0.306	0.120	0.014	-0.541	-0.071	-2.549	0.011
von Geusau, et al., 2004 men	-0.312	0.117	0.014	-0.541	-0.084	-2.676	0.007
von Geusau, et al., 2004 women	-0.347	0.103	0.011	-0.548	-0.145	-3.374	0.001
Wareing, et al., 2000	-0.282	0.116	0.013	-0.509	-0.055	-2.435	0.015
Yip & Lee, 2005	-0.285	0.123	0.015	-0.527	-0.044	-2.315	0.021
Zakzanis, et al., 2002	-0.316	0.118	0.014	-0.546	-0.085	-2.683	0.007
Overall ES	-0.297	0.113	0.013	-0.518	-0.075	-2.625	0.009

Information listed on row of named study represents the aggregate estimate with that study removed.

Table 12: Aggregate estimated effect size for measures of immediate verbal memory using Hedge's *g* and the random effects model with one study removed.

Study	ES	Standard error	Variance	Lower limit	Upper limit	*Z*-Value	*p*-Value
Back-Madruga, et al., 2003	-0.679	0.106	0.011	-0.887	-0.472	-6.422	0.000
Bhattachary & Powell, 2001	-0.601	0.112	0.013	-0.820	-0.382	-5.374	0.000
Croft, et al., 2001	-0.636	0.116	0.013	-0.863	-0.408	-5.470	0.000
Curran & Verheyden, 2003	-0.642	0.116	0.013	-0.869	-0.415	-5.538	0.000
Daumann, et al., 2003	-0.654	0.114	0.013	-0.878	-0.431	-5.744	0.000
Fisk, et al., 2004	-0.635	0.119	0.014	-0.869	-0.401	-5.321	0.000
Fox, Parrott, et al., 2001	-0.650	0.115	0.013	-0.875	-0.425	-5.657	0.000
Gouzoulis-Mayfrank, et al., 2000	-0.646	0.117	0.014	-0.875	-0.417	-5.529	0.000
Gouzoulis-Mayfrank, et al., 2003	-0.648	0.117	0.014	-0.877	-0.419	-5.550	0.000
Halpern, et al., 2004	-0.646	0.116	0.013	-0.873	-0.418	-5.567	0.000
Hanson & Luciana, 2004	-0.654	0.115	0.013	-0.880	-0.428	-5.670	0.000
McCardle, et al., 2004	-0.642	0.116	0.013	-0.870	-0.415	-5.542	0.000
Montgomery, et al., 2005	-0.630	0.117	0.014	-0.860	-0.401	-5.394	0.000
Morgan, 1998 study 2, 1999	-0.621	0.116	0.013	-0.849	-0.394	-5.349	0.000
Parrott, et al., 1998	-0.619	0.114	0.013	-0.843	-0.395	-5.410	0.000
Reneman, Majoie, et al., 2001	-0.626	0.115	0.013	-0.851	-0.401	-5.460	0.000
Reneman, Lavalaye, et al., 2001	-0.616	0.115	0.013	-0.841	-0.392	-5.375	0.000
Rodgers, 2000	-0.643	0.116	0.013	-0.870	-0.416	-5.557	0.000
Simon & Mattick, 2002	-0.645	0.118	0.014	-0.876	-0.414	-5.473	0.000
Verkes, et al., 2001	-0.623	0.116	0.013	-0.851	-0.396	-5.368	0.000
von Geusau, et al., 2004 men	-0.584	0.106	0.011	-0.791	-0.376	-5.519	0.000
von Geusau, et al., 2004 women	-0.656	0.114	0.013	-0.879	-0.433	-5.764	0.000
Wareing, et al., 2005	-0.610	0.114	0.013	-0.833	-0.387	-5.353	0.000
Wareing, Fisk, et al., 2004	-0.643	0.116	0.014	-0.871	-0.414	-5.517	0.000
Yip & Lee, 2005	-0.571	0.097	0.009	-0.762	-0.380	-5.867	0.000
Zakzanis, et al., 2003	-0.646	0.116	0.013	-0.873	-0.419	-5.586	0.000
Overall ES	-0.634	0.112	0.013	-0.853	-0.414	-5.653	0.000

Information listed on row of named study represents the aggregate estimate with that study removed.

Table 13: Aggregate estimated effect size for measures of delayed verbal memory using Hedge's *g* and the random effects model with one study removed.

Study	ES	Standard error	Variance	Lower limit	Upper limit	*Z*-Value	*p*-Value
Back-Madruga, et al., 2003	-0.954	0.329	0.108	-1.599	-0.309	-2.900	0.004
Bhattachary & Powell, 2001	-0.811	0.328	0.107	-1.453	-0.169	-2.477	0.013
Croft, et al., 2001	-0.907	0.335	0.112	-1.563	-0.250	-2.705	0.007
Curran & Verheyden, 2003	-0.897	0.335	0.113	-1.554	-0.239	-2.673	0.008
Fox, Parrott, et al., 2001	-0.951	0.329	0.108	-1.595	-0.307	-2.893	0.004
Gouzoulis-Mayfrank, et al., 2000	-0.908	0.341	0.116	-1.576	-0.240	-2.664	0.008
Gouzoulis-Mayfrank, et al., 2003	-0.888	0.342	0.117	-1.560	-0.217	-2.594	0.009
Halpern, et al., 2004	-0.914	0.336	0.113	-1.572	-0.255	-2.720	0.007
McCardle, et al., 2004	-0.910	0.335	0.112	-1.566	-0.253	-2.716	0.007
Morgan, 1998 study 2, 1999	-0.884	0.339	0.115	-1.549	-0.219	-2.606	0.009
Parrott, et al., 1998	-0.865	0.331	0.110	-1.515	-0.216	-2.611	0.009
Reneman, et al., 2000	-0.843	0.328	0.108	-1.486	-0.200	-2.571	0.010
Reneman, Majoie, et al., 2001	-0.870	0.330	0.109	-1.518	-0.223	-2.636	0.008
Reneman, Lavalaye, et al., 2001	-0.866	0.334	0.111	-1.520	-0.212	-2.595	0.009
Rodgers, 2000	-0.892	0.335	0.112	-1.548	-0.236	-2.666	0.008
Simon & Mattick, 2002	-0.917	0.344	0.118	-1.590	-0.243	-2.668	0.008
Yip & Lee, 2005	-0.575	0.141	0.020	-0.851	-0.299	-4.089	0.000
Zakzanis, et al., 2003	-0.945	0.330	0.109	-1.592	-0.298	-2.862	0.004
Overall ES	-0.880	0.317	0.101	-1.502	-0.258	-2.773	0.006

Information listed on row of named study represents the aggregate estimate with that study removed.

Table 14: Aggregate estimated effect size for measures of immediate nonverbal memory using Hedge's *g* and the random effects model with one study removed.

Study	ES	Standard error	Variance	Lower limit	Upper limit	Z-Value	*p*-Value
Back-Madruga, et al., 2003	-0.451	0.098	0.010	-0.643	-0.259	-4.598	0.000
Croft, et al., 2001	-0.437	0.099	0.010	-0.632	-0.243	-4.412	0.000
Daumann, et al., 2004	-0.413	0.097	0.009	-0.604	-0.222	-4.234	0.000
Fox, et al., 2002	-0.430	0.100	0.010	-0.626	-0.234	-4.306	0.000
Fox, Parrott, et al., 2001	-0.398	0.094	0.009	-0.582	-0.214	-4.240	0.000
Gouzoulis-Mayfrank, et al., 2000	-0.423	0.101	0.010	-0.620	-0.225	-4.201	0.000
Gouzoulis-Mayfrank, et al., 2003	-0.423	0.101	0.010	-0.620	-0.225	-4.193	0.000
Halpern, et al., 2004	-0.425	0.100	0.010	-0.621	-0.230	-4.264	0.000
Hanson & Luciana, 2004	-0.453	0.098	0.010	-0.644	-0.262	-4.645	0.000
McCann, et al., 1999	-0.456	0.096	0.009	-0.643	-0.268	-4.768	0.000
Morgan, 1998 study 1	-0.442	0.098	0.010	-0.634	-0.249	-4.499	0.000
Rodgers, 2000	-0.448	0.097	0.009	-0.638	-0.257	-4.605	0.000
Simon & Mattick, 2002	-0.462	0.095	0.009	-0.649	-0.275	-4.844	0.000
Verkes, et al., 2001	-0.401	0.096	0.009	-0.589	-0.213	-4.176	0.000
von Geusau, et al., 2004 men	-0.385	0.087	0.008	-0.554	-0.215	-4.437	0.000
von Geusau, et al., 2004 women	-0.446	0.097	0.009	-0.637	-0.255	-4.585	0.000
Wareing, et al., 2005	-0.405	0.097	0.009	-0.595	-0.216	-4.194	0.000
Wareing, Murphy, et al., 2004	-0.412	0.097	0.009	-0.603	-0.221	-4.236	0.000
Yip & Lee, 2005	-0.413	0.103	0.011	-0.615	-0.211	-4.011	0.000
Zakzanis, et al., 2003	-0.449	0.097	0.009	-0.639	-0.258	-4.611	0.000
Overall ES	-0.428	0.095	0.009	-0.614	-0.242	-4.511	0.000

Information listed on row of named study represents the aggregate estimate with that study removed.

Table 15: Aggregate estimated effect size for measures of delayed nonverbal memory using Hedge's *g* and the random effects model with one study removed.

Study	ES	Standard error	Variance	Lower limit	Upper limit	Z-Value	*p*-Value
Back-Madruga, et al., 2003	-0.691	0.291	0.085	-1.261	-0.121	-2.375	0.018
Bhattachary & Powell, 2001	-0.702	0.285	0.081	-1.261	-0.142	-2.459	0.014
Croft, et al., 2001	-0.694	0.286	0.082	-1.254	-0.133	-2.424	0.015
Gouzoulis-Mayfrank, et al., 2003	-0.634	0.300	0.090	-1.222	-0.045	-2.110	0.035
Halpern, et al., 2004	-0.669	0.291	0.085	-1.240	-0.098	-2.297	0.022
McCann, et al., 1999	-0.372	0.202	0.041	-0.769	0.024	-1.839	0.066
Rodgers, 2000	-0.654	0.290	0.084	-1.221	-0.086	-2.258	0.024
Simon & Mattick, 2002	-0.703	0.291	0.085	-1.274	-0.132	-2.412	0.016
Yip & Lee, 2005	-0.542	0.269	0.073	-1.070	-0.014	-2.012	0.044
Zakzanis, et al., 2003	-0.711	0.282	0.080	-1.265	-0.158	-2.521	0.012
Overall ES	-0.634	0.265	0.070	-1.154	-0.114	-2.390	0.017

Information listed on row of named study represents the aggregate estimate with that study removed.

Table 16: Aggregate estimated effect size for measures of executive functioning using Hedge's *g* and the random effects model with one study removed.

Study	ES	Standard error	Variance	Lower limit	Upper limit	Z-Value	*p*-Value
Back-Madruga, et al., 2003	-0.272	0.104	0.011	-0.476	-0.068	-2.608	0.009
Fisk, et al., 2004	-0.277	0.109	0.012	-0.490	-0.063	-2.543	0.011
Fox, Parrott, et al., 2001	-0.235	0.107	0.012	-0.446	-0.025	-2.192	0.028
Gouzoulis-Mayfrank, et al., 2000	-0.174	0.094	0.009	-0.358	0.010	-1.857	0.063
Gouzoulis-Mayfrank, et al., 2003	-0.242	0.111	0.012	-0.460	-0.024	-2.177	0.030
Halpern, et al., 2004	-0.215	0.105	0.011	-0.421	-0.009	-2.049	0.040
Montgomery, et al., 2005	-0.240	0.110	0.012	-0.455	-0.025	-2.189	0.029
Morgan, 1998 study 1	-0.244	0.108	0.012	-0.455	-0.033	-2.267	0.023
Morgan, 1998 study 2, 1999	-0.269	0.104	0.011	-0.474	-0.064	-2.576	0.010
von Geusau, et al., 2004 men	-0.227	0.103	0.011	-0.430	-0.025	-2.198	0.028
von Geusau, et al., 2004 women	-0.265	0.091	0.008	-0.443	-0.086	-2.911	0.004
Wareing, et al., 2000	-0.203	0.095	0.009	-0.390	-0.017	-2.135	0.033
Zakzanis, et al., 2003	-0.218	0.105	0.011	-0.423	-0.012	-2.078	0.038
Overall ES	-0.238	0.100	0.010	-0.433	-0.042	-2.385	0.017

Information listed on row of named study represents the aggregate estimate with that study removed.

Table 17: Aggregate estimated effect size for measures of fluency using Hedge's *g* and the random effects model with one study removed.

Study	ES	Standard error	Variance	Lower limit	Upper limit	Z-Value	*p*-Value
Back-Madruga, et al., 2003	-0.541	0.119	0.014	-0.775	-0.308	-4.541	0.000
Bhattachary & Powell, 2001	-0.370	0.122	0.015	-0.608	-0.131	-3.040	0.002
Croft, et al., 2001	-0.426	0.151	0.023	-0.723	-0.130	-2.823	0.005
Curran & Verheyden, 2003	-0.486	0.150	0.022	-0.779	-0.193	-3.248	0.001
Fox, et al., 2002	-0.455	0.155	0.024	-0.760	-0.151	-2.936	0.003
Gouzoulis-Mayfrank, et al., 2000	-0.466	0.157	0.025	-0.775	-0.158	-2.963	0.003
Halpern, et al., 2004	-0.460	0.155	0.024	-0.763	-0.158	-2.980	0.003
Hanson & Luciana, 2004	-0.485	0.153	0.024	-0.786	-0.184	-3.161	0.002
Heffernan, et al., 2001 study 2	-0.400	0.148	0.022	-0.690	-0.109	-2.697	0.007
Yip & Lee, 2005	-0.439	0.168	0.028	-0.768	-0.110	-2.615	0.009
Overall ES	-0.452	0.140	0.020	-0.727	-0.177	-3.220	0.001

Information listed on row of named study represents the aggregate estimate with that study removed.

Table 18: Aggregate estimated effect size for measures of reaction time using Hedge's *g* and the random effects model with one study removed.

Study	ES	Standard error	Variance	Lower limit	Upper limit	Z-Value	*p*-Value
Daumann, et al., 2003	-0.446	0.111	0.012	-0.663	-0.229	-4.025	0.000
Fox, et al., 2002	-0.441	0.114	0.013	-0.665	-0.218	-3.874	0.000
Fox, Parrott, et al., 2001	-0.430	0.112	0.013	-0.649	-0.211	-3.845	0.000
Gouzoulis-Mayfrank, et al., 2000	-0.450	0.117	0.014	-0.679	-0.221	-3.852	0.000
Gouzoulis-Mayfrank, et al., 2003	-0.454	0.118	0.014	-0.684	-0.223	-3.849	0.000
Jacobsen, et al., 2004	-0.402	0.109	0.012	-0.615	-0.189	-3.695	0.000
Parrott, et al., 1998	-0.468	0.111	0.012	-0.685	-0.252	-4.236	0.000
Rodgers, 2000	-0.456	0.112	0.013	-0.676	-0.236	-4.067	0.000
Verkes, et al., 2001	-0.394	0.114	0.013	-0.617	-0.171	-3.463	0.001
von Geusau, et al., 2004 men	-0.338	0.109	0.012	-0.552	-0.124	-3.091	0.002
von Geusau, et al., 2004 women	-0.472	0.111	0.012	-0.689	-0.255	-4.260	0.000
Overall ES	-0.431	0.107	0.011	-0.641	-0.221	-4.027	0.000

Information listed on row of named study represents the aggregate estimate with that study removed.

Table 19: Aggregate estimated effect size for measures of learning using Hedge's *g* and the random effects model with one study removed.

Study	ES	Standard error	Variance	Lower limit	Upper limit	Z-Value	*p*-Value
Back-Madruga, et al., 2003	-0.243	0.122	0.015	-0.482	-0.004	-1.989	0.047
Croft, et al., 2001	-0.110	0.118	0.014	-0.342	0.122	-0.928	0.353
Curran & Verheyden, 2003	-0.138	0.118	0.014	-0.371	0.094	-1.169	0.242
Fox, et al., 2002	-0.154	0.120	0.014	-0.390	0.081	-1.284	0.199
Gouzoulis-Mayfrank, et al., 2000	-0.148	0.123	0.015	-0.390	0.094	-1.196	0.232
Halpern, et al., 2004	-0.153	0.120	0.014	-0.388	0.082	-1.278	0.201
McCardle, et al., 2004	-0.138	0.118	0.014	-0.370	0.094	-1.167	0.243
Zakzanis, et al., 2003	-0.131	0.118	0.014	-0.363	0.101	-1.107	0.268
Overall ES	-0.151	0.112	0.013	-0.371	0.068	-1.349	0.177

Information listed on row of named study represents the aggregate estimate with that study removed.

APPENDIX B
FIGURES

Figure 1: Forrest plot of primary study effect sizes for combined cognitive domains.

Model	Study name	Outcome	Hedges's g and 95% CI
	Back-Madruga, et al., 2003	Combined	
	Bhattachary & Powell, 2001	Combined	
	Bond, et al., 2004	processing speed	
	Croft, et al., 2001	Combined	
	Curran & Verheyden, 2003	Combined	
	Daumann, et al., 2003	Combined	
	Daumann, et al., 2004	LGT-3 Logos immediate recall	
	Fisk, et al., 2004	Combined	
	Fox, et al., 2002	Combined	
	Fox, Parrott, et al., 2001	Combined	
	Gouzoulis-Mayfrank, et al., 2000	Combined	
	Gouzoulis-Mayfrank, et al., 2003	Combined	
	Halpern, et al., 2004	Combined	
	Hanson & Luciana, 2004	Combined	
	Heffernan, et al., 2001 study 2	Combined	
	Jacobsen, et al., 2004	Combined	
	McCann, et al., 1999	Combined	
	McCardle, et al., 2004	Combined	
	Montgomery, et al., 2005	Combined	
	Morgan, 1998 study 1	Combined	
	Morgan, 1998 study 2, 1999	Combined	
	Parrott, et al., 1998	Combined	
	Reneman, et al., 2000	Rey Auditory Verbal Learning Test delayed recall	
	Reneman, Lavalaye, et al., 2001	Combined	
	Reneman, Majoie, et al., 2001	Combined	
	Rodgers, 2000	Combined	
	Simon & Mattick, 2002	Combined	
	Verkes, et al., 2001	Combined	
	von Geusau, et al., 2004 men	Combined	
	von Geusau, et al., 2004 women	Combined	
	Wareing, et al., 2000	Combined	
	Wareing, et al., 2005	Combined	
	Wareing, Fisk, et al., 2004	Combined	
	Wareing, Murphy, et al., 2004	Combined	
	Yip & Lee, 2005	Combined	
	Zakzanis, et al., 2002	Combined	
	Zakzanis, et al., 2003	Combined	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship

81

Figure 2: Funnel plot for the aggregate effect size for combined cognitive domains under the random effects model.

Figure 3: Forrest plot for measures of attention.

Model	Study name	Outcome	Hedges's g and 95% CI
	Back-Madruga, et al., 2003	Combined	
	Bond, et al., 2004	processing speed	
	Croft, et al., 2001	Combined	
	Curran & Verheyden, 2003	Combined	
	Fox, et al., 2002	Combined	
	Gouzoulis-Mayfrank, et al., 2000	Combined	
	Gouzoulis-Mayfrank, et al., 2003	Combined	
	Halpern, et al., 2004	Combined	
	Hanson & Luciana, 2004	Combined	
	Jacobsen, et al., 2004	Combined	
	McCann, et al., 1999	Combined	
	McCardle, et al., 2004	Combined	
	Morgan, 1998 study 1	Combined	
	Morgan, 1998 study 2, 1999	Combined	
	Parrott, et al., 1998	Combined	
	Rodgers, 2000	WMS-III mental control	
	Simon & Mattick, 2002	WMS-III working	
	von Geusau, et al., 2004 men	Combined	
	von Geusau, et al., 2004 women	Combined	
	Wareing, et al., 2000	Combined	
	Yip & Lee, 2005	Combined	
	Zakzanis, et al., 2002	Combined	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship

Figure 4: Forrest plot for measures of immediate verbal memory

Model	Studyname	Outcome	Hedges's g and 95% CI
	Back-Madruga, et al., 2003	Combined	
	Bhattachary & Powell, 2001	Combined	
	Croft, et al., 2001	Combined	
	Curran & Verheyden, 2003	Combined	
	Daumann, et al., 2003	Combined	
	Fisk, et al., 2004	Computation Span	
	Fox, Parrott, et al., 2001	Combined	
	Gouzoulis-Mayfrank, et al., 2000	Combined	
	Gouzoulis-Mayfrank, et al., 2003	Combined	
	Halpern, et al., 2004	Combined	
	Hanson & Luciana, 2004	Combined	
	McCardle, et al., 2004	Combined	
	Montgomery, et al., 2005	Combined	
	Morgan, 1998 study 2, 1999	Rivermead Behavioral Memory Test immediate prose recall	
	Parrott, et al., 1998	immediate word recall	
	Reneman, Lavalaye, et al., 2001	Rey Auditory Verbal Learning Test immediate recall	
	Reneman, Majoie, et al., 2001	Rey Auditory Verbal Learning Test immediate recall	
	Rodgers, 2000	Combined	
	Simon & Mattick, 2002	WMS-III auditory immediate	
	Verkes, et al., 2001	Combined	
	von Geusau, et al., 2004 men	Mental Counters % correct	
	von Geusau, et al., 2004 women	Mental Counters % correct	
	Wareing, et al., 2005	computation span	
	Wareing, Fisk, et al., 2004	Combined	
	Yip & Lee, 2005	Combined	
	Zakzanis, et al., 2003	Rivermead Behavioral Memory Test immediate prose recall	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship

84

Figure 5: Forrest plot for measures of delayed verbal memory

Model	Studyname	Outcome	Hedges's g and 95% CI
	Back-Madruga, et al., 2003	Combined	
	Bhattachary & Powell, 2001	prose recall delayed	
	Croft, et al., 2001	Coughlan list B	
	Curran & Verheyden, 2003	Combined	
	Fox, Parrott, et al., 2001	delayed prose recall	
	Gouzoulis-Mayfrank, et al., 2000	VLMT delayed recall	
	Gouzoulis-Mayfrank, et al., 2003	LGT-3 German-Turkish delayed recall	
	Halpern, et al., 2004	Combined	
	McCardle, et al., 2004	Rey Auditory Verbal Learning Test trial 8	
	Morgan, 1998 study 2, 1999	Rivermead Behavioral Memory Test delayed prose recall	
	Parrott, et al., 1998	delayed word recall	
	Reneman, et al., 2000	Rey Auditory Verbal Learning Test delayed recall	
	Reneman, Lavalaye, et al., 2001	Rey Auditory Verbal Learning Test delayed recall	
	Reneman, Majoie, et al., 2001	Rey Auditory Verbal Learning Test delayed recall	
	Rodgers, 2000	Combined	
	Simon & Mattick, 2002	Combined	
	Yip & Lee, 2005	Combined	
	Zakzanis, et al., 2003	Rivermead Behavioral Memory Test delayed prose recall	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship

Figure 6: Forrest plot for measures of immediate nonverbal memory

Model	Study name	Outcome	Hedges's g and 95% CI
	Back-Madruga, et al., 2003	Combined	
	Croft, et al., 2001	Coughlan design 6	
	Daumann, et al., 2004	LGT-3 Logos immediate recall	
	Fox, et al., 2002	Combined	
	Fox, Parrott, et al., 2001	spatial recall	
	Gouzoulis-Mayfrank, et al., 2000	Combined	
	Gouzoulis-Mayfrank, et al., 2003	Combined	
	Halpern, et al., 2004	Combined	
	Hanson & Luciana, 2004	Combined	
	McCann, et al., 1999	Combined	
	Morgan, 1998 study 1	Combined	
	Rodgers, 2000	Combined	
	Simon & Mattick, 2002	WMS-III visual immediate	
	Verkes, et al., 2001	Combined	
	von Geusau, et al., 2004 men	Tic Tac Toe % correct	
	von Geusau, et al., 2004 women	Tic Tac Toe % correct	
	Wareing, et al., 2005	Combined	
	Wareing, Murphy, et al., 2004	Combined	
	Yip & Lee, 2005	Aggie Figures Learning Test immediate	
	Zakzanis, et al., 2003	Rivermead Behavioral Memory Test immediate route	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship

86

Figure 7: Forrest plot for measures of delayed nonverbal memory

Model	Study name	Outcome	Hedges's g and 95% CI
	Back-Madruga, et al., 2003	Combined	
	Bhattachary & Powell, 2001	Rey-Osterreith recall	
	Croft, et al., 2001	Coughlan design B	
	Gouzoulis-Mayfrank, et al., 2003	Combined	
	Halpern, et al., 2004	Combined	
	McCann, et al., 1999	delayed code substitution with no key	
	Rodgers, 2000	Combined	
	Simon & Mattick, 2002	WMS-III visual delayed	
	Yip & Lee, 2005	Combined	
	Zakzanis, et al., 2003	Combined	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship

Figure 8: Forrest plot for measures of executive functioning.

Model	Study name	Outcome	Hedges's g and 95%CI
	Back-Madruga, et al., 2003	Combined	
	Fisk, et al., 2004	Combined	
	Fox, Parrott, et al., 2001	Combined	
	Gouzoulis-Mayfrank, et al., 2000	Combined	
	Gouzoulis-Mayfrank, et al., 2003	Combined	
	Halpern, et al., 2004	Combined	
	Montgomery, et al., 2005	Combined	
	Morgan, 1998 study 1	Combined	
	Morgan, 1998 study 2, 1999	Combined	
	von Geusau, et al., 2004 men	Combined	
	von Geusau, et al., 2004 women	Combined	
	Wareing, et al., 2000	Combined	
	Zakzanis, et al., 2003	Combined	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship

Figure 9: Forrest plot for measures of fluency.

Model	Study name	Outcome	Hedges's g and 95%CI
	Back-Madruga, et al., 2003	Combined	
	Bhattachary & Powell, 2001	verbal fluency	
	Croft, et al., 2001	Combined	
	Curran & Verheyden, 2003	phonemic fluency	
	Fox, et al., 2002	Combined	
	Gouzoulis-Mayfrank, et al., 2000	Combined	
	Halpern, et al., 2004	Combined	
	Hanson & Luciana, 2004	Combined	
	Heffernan, et al., 2001 study 2	Combined	
	Yip & Lee, 2005	Combined	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship

Figure 10: Forrest plot for measures of reaction time.

Model	Study name	Outcome	Hedges's g and 95% CI
	Daumann, et al., 2003	Combined	
	Fox, et al., 2002	Combined	
	Fox, Parrott, et al., 2001	simple reaction time	
	Gouzoulis-Mayfrank, et al., 2000	Combined	
	Gouzoulis-Mayfrank, et al., 2003	Combined	
	Jacobsen, et al., 2004	Combined	
	Parrott, et al., 1998	simple reaction time	
	Rodgers, 2000	Combined	
	Verkes, et al., 2001	Combined	
	von Geusau, et al., 2004 men	Combined	
	von Geusau, et al., 2004 women	Combined	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship

Figure 11: Forrest plot for measures of learning.

Model	Studyname	Outcome	Hedges's g and 95% CI
	Back-Madruga, et al., 2003	Rey Auditory Verbal Learning Test trial 5	
	Croft, et al., 2001	Combined	
	Curran & Verheyden, 2003	Buschke SRT 3-1	
	Fox, et al., 2002	Combined	
	Gouzoulis-Mayfrank, et al., 2000	Combined	
	Halpern, et al., 2004	Combined	
	McCardle, et al., 2004	Combined	
	Zakzanis, et al., 2003	Rivermead Behavioral Memory Test 1st/2nd name	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship

91

Figure 12: Forrest plot of primary studies that used non-substance users as the control group. Cognitive domains were combined

Figure 13: Forrest plot of primary studies that used a mixed group of substance users and non-substance users as the control group. Cognitive domains were combined.

Model	Study name	Outcome	Hedges's g and 95% CI
	Bond, et al., 2004	processing speed	
	Fisk, et al., 2004	Combined	
	Fox, Parrott, et al., 2001	Combined	
	Halpern, et al., 2004	Combined	
	Heffernan, et al., 2001 study 2	Combined	
	Jacobsen, et al., 2004	Combined	
	McCann, et al., 1999	Combined	
	McCardle, et al., 2004	Combined	
	Montgomery, et al., 2005	Combined	
	Reneman, et al., 2001	Combined	
	Reneman, Lavalaye, et al., 2001	Combined	
	Verkes, et al., 2001	Combined	
	von Geusau, et al., 2004 men	Combined	
	von Geusau, et al., 2004 women	Combined	
	Wareing, et al., 2005	Combined	
	Wareing, Fisk, et al., 2004	Combined	
	Wareing, Murphy, et al., 2004	Combined	
	Zakzanis, et al., 2002	Combined	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship

93

Figure 14: Forrest plot of primary studies that used marijuana users as the control group. Cognitive domains were combined

Model	Study name	Outcome	Hedges's g and 95% CI
	Curran & Verheyden, 2003	Combined	
	Daumann, et al., 2004	LGT-3 Logos immediate recall	
	Gouzoulis-Mayfrank, et al., 2000	Combined	
	Gouzoulis-Mayfrank, et al., 2003	Combined	
	Rodgers, 2000	Combined	
	Simon & Mattick, 2002	Combined	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship
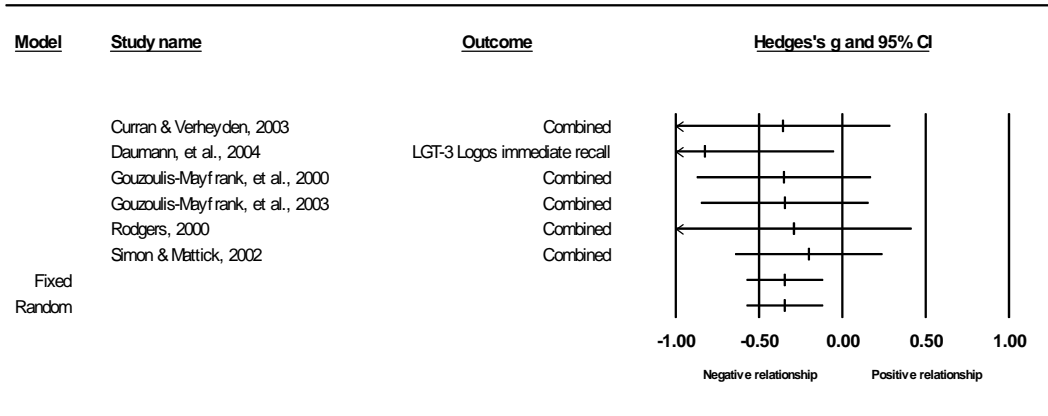
Figure 15: Forrest plot of primary studies that used polysubstance users as the control group. Cognitive domains were combined.

Model	Study name	Outcome	Hedges's g and 95% CI
	Fox, et al., 2002	Combined	
	Morgan, 1998 study 1	Combined	
	Morgan, 1998 study 2, 1999	Combined	
	Zakzanis, et al., 2003	Combined	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship
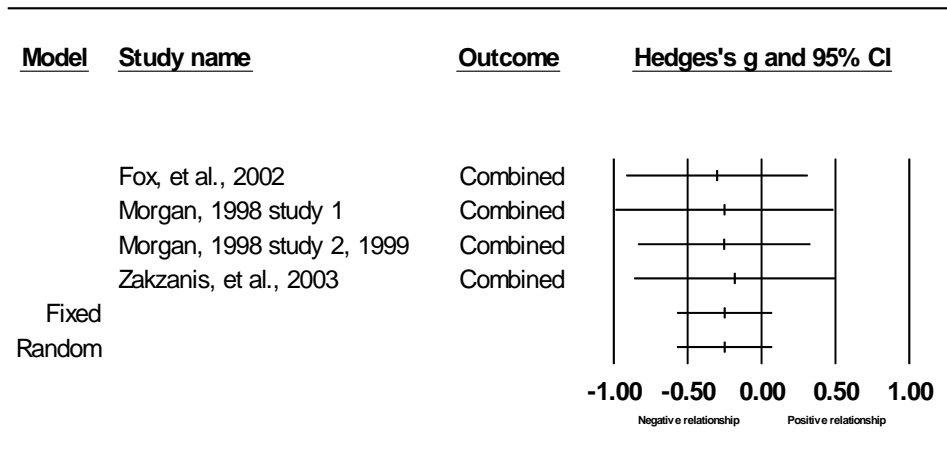
Figure 16: Mixed effects meta-regression using unrestricted maximum likelihood to test reported total lifetime cumulative ecstasy use as a potential moderator.

Regression of lifetime amount MDMA on Hedges's g
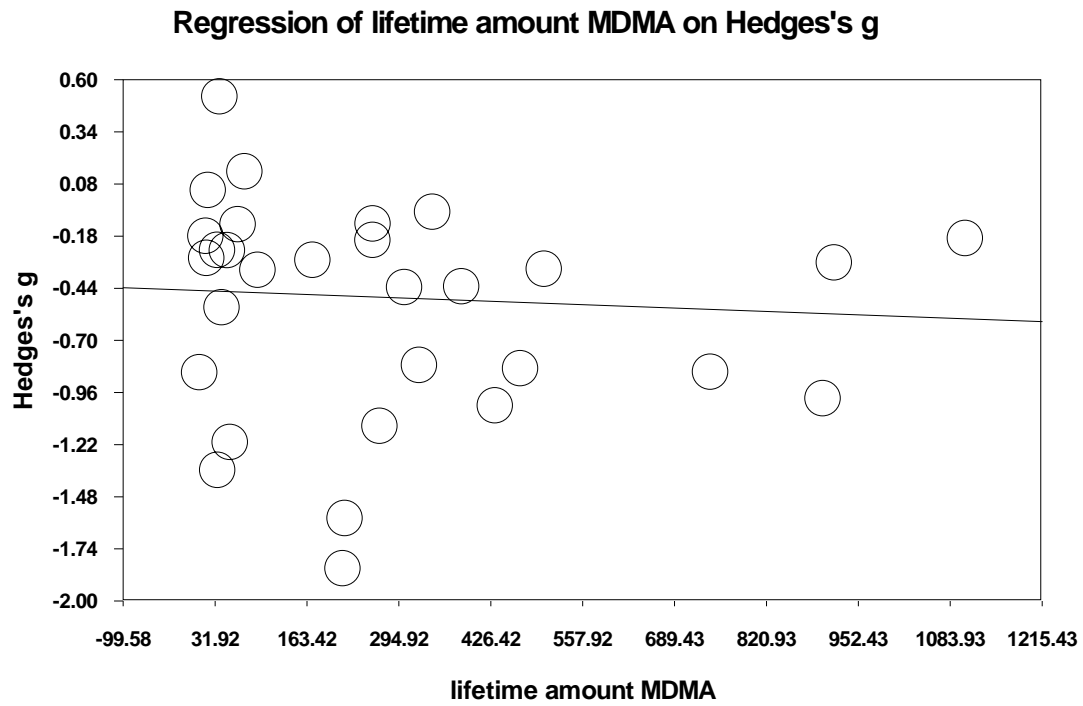
Figure 17: Mixed effects meta-regression using unrestricted maximum likelihood to test reported frequency of ecstasy use per month as a potential moderator.

Regression of frequency MDMA used per month on Hedges's g
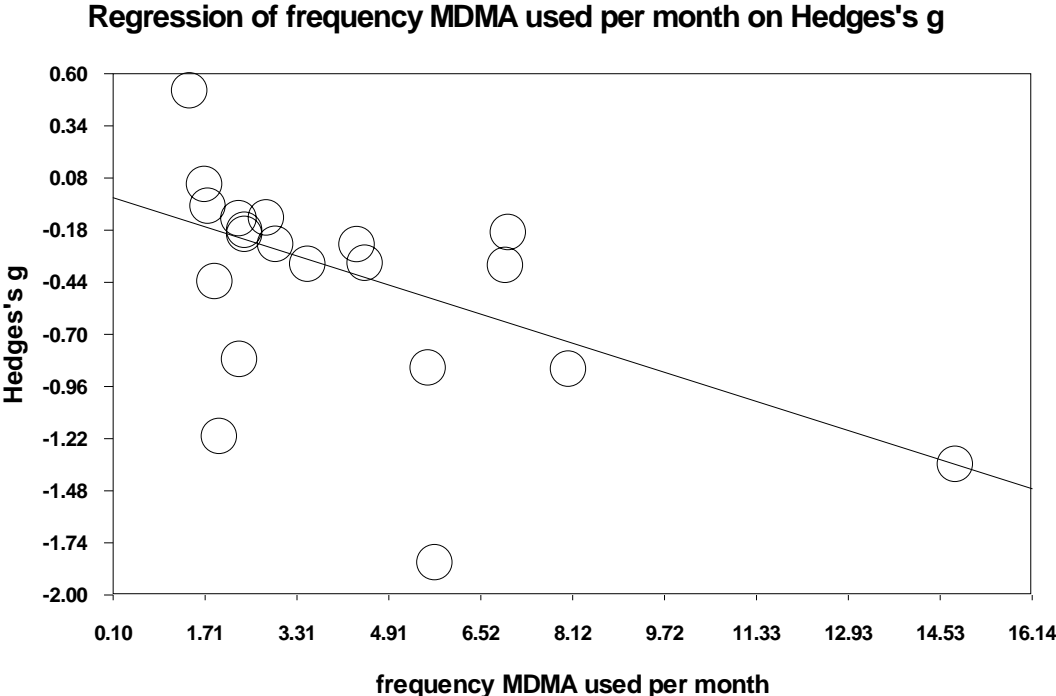
Figure 18: Mixed effects meta-regression using unrestricted maximum likelihood to test reported number of years participants used ecstasy as a potential moderator.

Regression of years used MDMA on Hedges's g
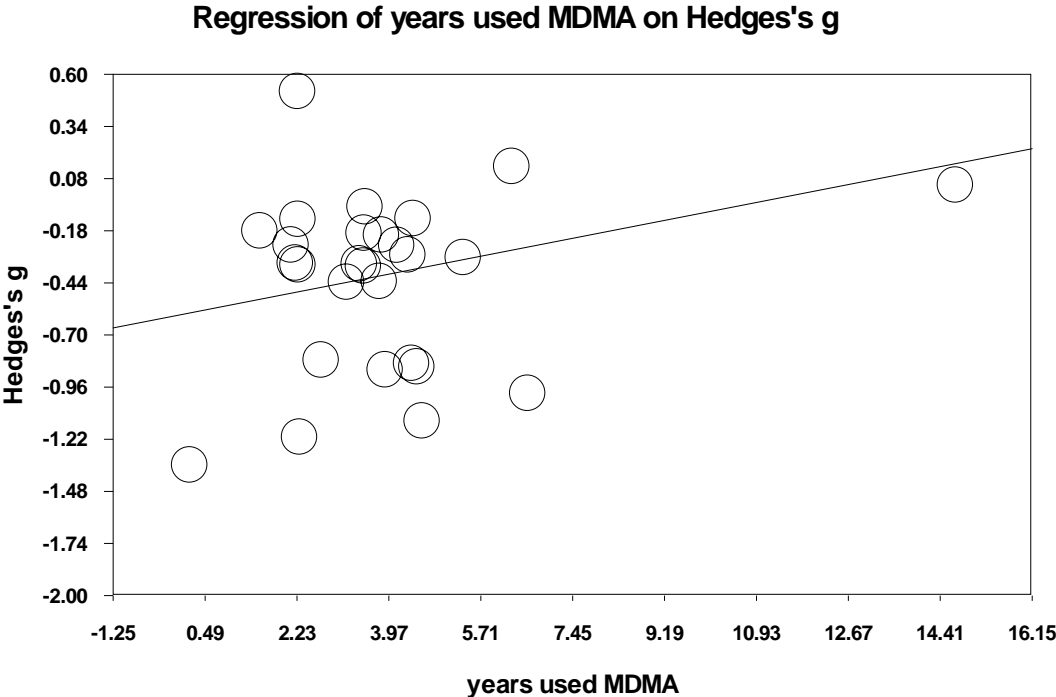
Figure 19: Mixed effects meta-regression using unrestricted maximum likelihood to test reported number of years participants used ecstasy as a potential moderator, with an extreme score removed.
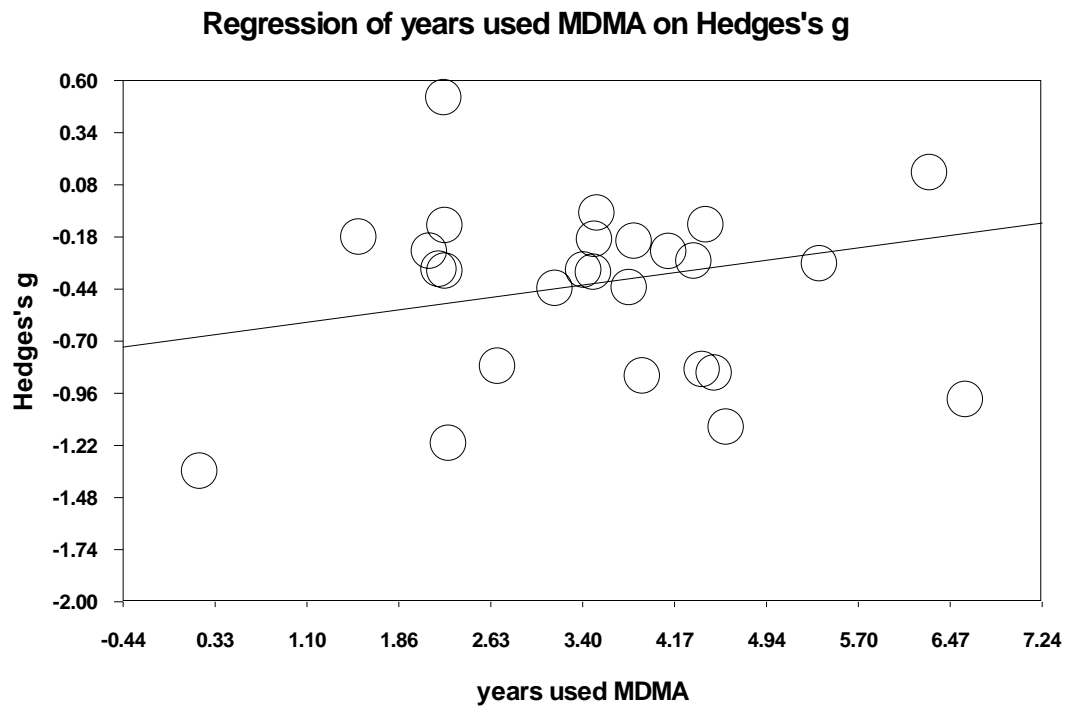
Regression of years used MDMA on Hedges's g

Figure 20: Mixed effects meta-regression using unrestricted maximum likelihood to test minimum required duration of abstinence before cognitive testing as a potential moderator.
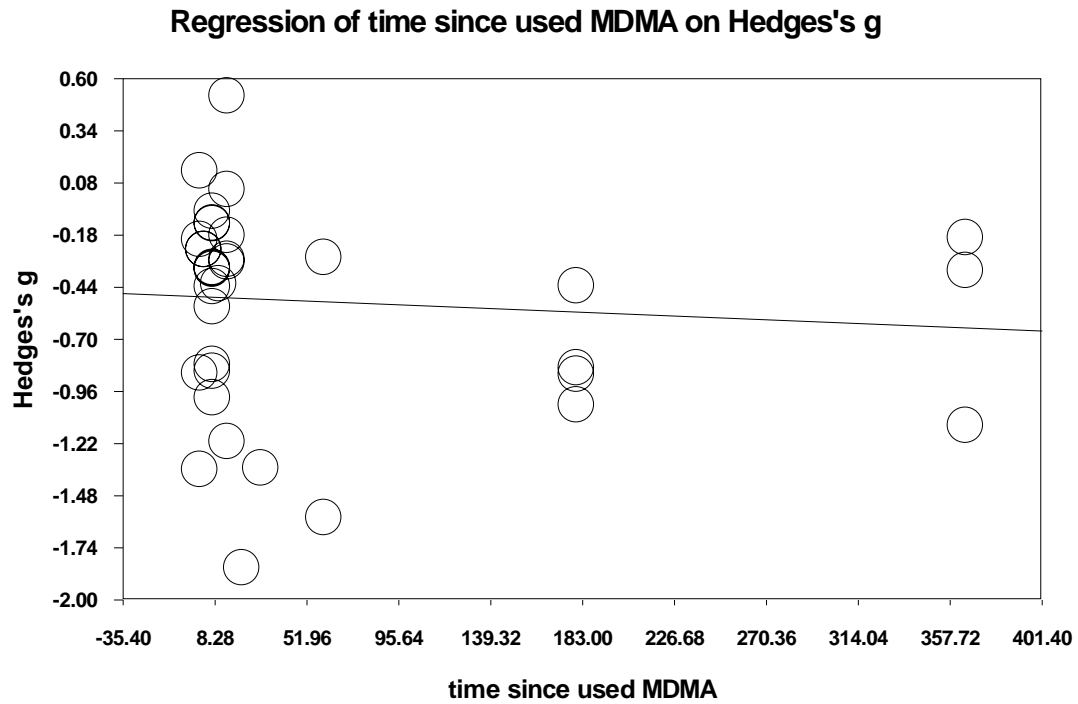
Regression of time since used MDMA on Hedges's g

Figure 21: Forrest plot for studies in which the control group was balanced for alcohol, marijuana, and other substance use.

Model	Study name	Outcome	Hedges's g and 95% CI
	Jacobsen, et al., 2004	Combined	
	Morgan, 1998 study 1	Combined	
	Morgan, 1998 study 2, 1999	Combined	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship
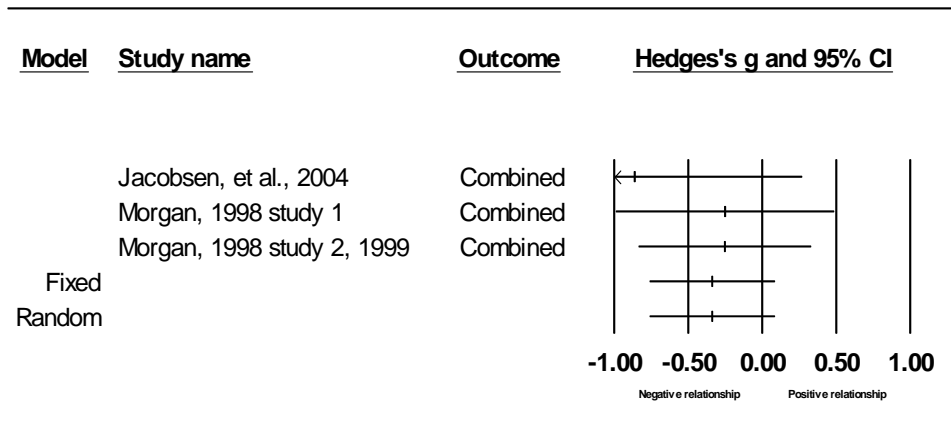
Figure 22: Forrest plot for studies in which the control group was balanced for intelligence, education, and age.

Model	Study name	Outcome	Hedges's g and 95% CI
	Back-Madruga, et al., 2003	Combined	
	Fisk, et al., 2004	Combined	
	Hanson & Luciana, 2004	Combined	
	Jacobsen, et al., 2004	Combined	
	McCardle, et al., 2004	Combined	
	Montgomery, et al., 2005	Combined	
	Morgan, 1998 study 1	Combined	
	Morgan, 1998 study 2, 1999	Combined	
	Wareing, Fisk, et al., 2004	Combined	
	Yip & Lee, 2005	Combined	
	Zakzanis, et al., 2002	Combined	
Fixed			
Random			

-1.00   -0.50   0.00   0.50   1.00

Negative relationship   Positive relationship

Figure 23: Aggregated estimated effect size using Hedge's *g* and the random effects model with one study removed.

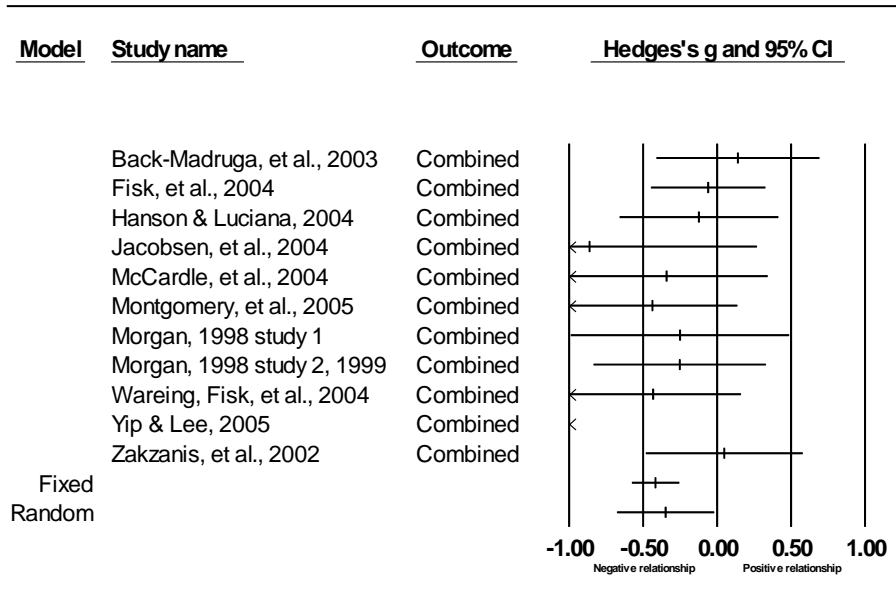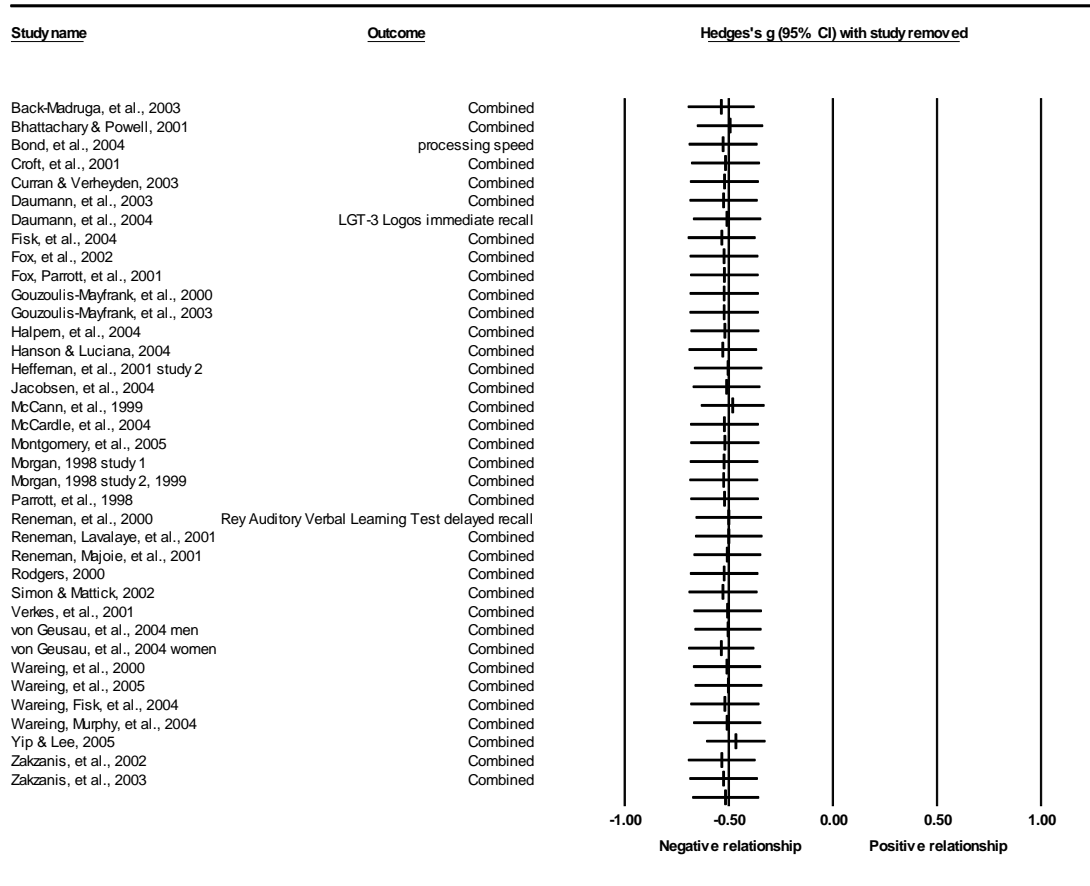| Study name	Outcome	Hedges's g (95% CI) with study removed
Back-Madruga, et al., 2003	Combined	
Bhattachary & Powell, 2001	Combined	
Bond, et al., 2004	processing speed	
Croft, et al., 2001	Combined	
Curran & Verheyden, 2003	Combined	
Daumann, et al., 2003	Combined	
Daumann, et al., 2004	LGT-3 Logos immediate recall	
Fisk, et al., 2004	Combined	
Fox, et al., 2002	Combined	
Fox, Parrott, et al., 2001	Combined	
Gouzoulis-Mayfrank, et al., 2000	Combined	
Gouzoulis-Mayfrank, et al., 2003	Combined	
Halpern, et al., 2004	Combined	
Hanson & Luciana, 2004	Combined	
Heffernan, et al., 2001 study 2	Combined	
Jacobsen, et al., 2004	Combined	
McCann, et al., 1999	Combined	
McCardle, et al., 2004	Combined	
Montgomery, et al., 2005	Combined	
Morgan, 1998 study 1	Combined	
Morgan, 1998 study 2, 1999	Combined	
Parrott, et al., 1998	Combined	
Reneman, et al., 2000	Rey Auditory Verbal Learning Test delayed recall	
Reneman, Lavalaye, et al., 2001	Combined	
Reneman, Majoie, et al., 2001	Combined	
Rodgers, 2000	Combined	
Simon & Mattick, 2002	Combined	
Verkes, et al., 2001	Combined	
von Geusau, et al., 2004 men	Combined	
von Geusau, et al., 2004 women	Combined	
Wareing, et al., 2000	Combined	
Wareing, et al., 2005	Combined	
Wareing, Fisk, et al., 2004	Combined	
Wareing, Murphy, et al., 2004	Combined	
Yip & Lee, 2005	Combined	
Zakzanis, et al., 2002	Combined	
Zakzanis, et al., 2003	Combined	

The last line of the Forrest plot represents the overall effect size estimate, computed under the random effects model.