Effects of Agonist-Antagonist Opioids in Humans Trained in a Hydromorphone/Not Hydromorphone Discrimination

KENZIE L. PRESTON and GEORGE E. BIGELOW

Behavioral Pharmacology Research Unit, Department of Psychiatry and Behavioral Sciences, The Johns Hopkins University School of Medicine, Baltimore, Maryland

Accepted for publication June 14, 2000 This paper is available online at http://www.jpet.org

ABSTRACT

The purpose of this study was to examine the discrimination of agonist-antagonist opioids in humans trained in a two-choice hydromorphone/not hydromorphone discrimination. Eight adult male volunteers with histories of opioid abuse who were not currently physically dependent were trained to discriminate the mu receptor agonist hydromorphone (3 mg/70 kg, i.m.) (“Drug A”) from a “Not Drug A” training condition (saline placebo). Volunteers received financial reinforcement for correct responses. After training, generalization dose-effect curves for hydromorphone, butorphanol, pentazocine, nalbuphine, and buprenorphine were determined. Other subjective, behavioral, and physiological measures were concurrently collected in all sessions. In generalization testing hydromorphone and buprenorphine produced dose-related increases in hydromorphone-appropriate responses. Pentazocine produced an inverted U-shaped dose-response curve with complete substitution at 32 mg/70 kg but not at 64 mg/70 kg. Butorphanol and nalbuphine did not completely substitute for hydromorphone at any dose tested. These results differ from an earlier two-choice, Drug A versus Drug B (hydromorphone/saline) discrimination study. After Drug/Not Drug instructions the behavioral discriminations of agonist-antagonist opioids were more consistent with their putative agonist activities at the mu opioid receptor and with their subjective effects profiles than was the case after Drug A versus Drug B instructions. These results suggest that instructions are an important factor in the outcome of human drug discrimination studies.

Drug discrimination has become a standard tool in the experimental laboratory for characterizing physiological and behavioral effects and neuropathological actions of drugs (Stolerman et al., 1995). The value of this procedure largely stems from its pharmacological specificity. That is, when animals trained to discriminate a drug from saline (or no drug) are tested with a variety of other drugs, they tend to emit drug-appropriate responses only in the presence of the training drug and other pharmacologically similar drugs and to emit an alternate response in the presence of saline and pharmacologically dissimilar drugs. The pharmacological specificity of drug discrimination responding has been extensively documented in animal studies (Young, 1991).

The pharmacological specificity of drug discrimination responding has been tested in humans, although to a lesser degree than in nonhumans (Kamien et al., 1993). For example, in stimulant abusers trained to discriminate d-amphetamine, 30 mg, and placebo, the stimulants d-amphetamine and methylphenidate fully substituted, and the sedative diazepam and opiate hydromorphone did not substitute, for d-amphetamine (Heishman and Henningfield, 1991; Lamb and Henningfield, 1994). Tests in normal volunteers trained to discriminate diazepam, 10 mg, from placebo have also been consistent with pharmacological specificity: diazepam, triazolam, and lorazepam fully substituted, pentobarbital and buspirone partially substituted, and d-amphetamine and tripelenamine did not substitute for diazepam (Johanson, 1991a,b). Thus, overall, there is cross generalization among stimulants, there is cross generalization among sedatives, and humans discriminate stimulants from sedatives, just as has been shown in the animal laboratory (Young, 1991).

Pharmacological specificity can extend to receptor subclasses. The fact that laboratory animals can discriminate among agonists acting at the mu and kappa opioid receptors (e.g., Herling and Woods, 1981; Picker and Cook, 1997) has made drug discrimination a particularly useful tool for studying opioid pharmacology. Conversely, receptor theory has been useful in developing a basis for interpreting opioid drug

ABBREVIATIONS: ARCI, Addiction Research Center Inventory; DSST, Digit Symbol Substitution Test; LSD, Lysergic Acid Diethylamide; MBG, Morphine-Benzedrine group; PCAG, Pentobarbital, Chlorpromazine, Alcohol group; BG, Benzedrine group; VAS, visual analog scale(s); CHO, Chinese hamster ovary; GTPγS, guanosine 5’-3-O-(thio)triphosphate; HSD, honestly significant difference.
discrimination results (Woods et al., 1988). Our laboratory has conducted a series of studies investigating the discriminative stimulus effects of a group of four agonist-antagonist opioids under a variety of training conditions in experienced opioid abusers (Preston et al., 1989, 1992; Preston and Bigelow, 1990, 1994; Jones et al., 1999). Butorphanol, pentazocine, nalbuphine, and buprenorphine have varying degrees of agonist and antagonist activity at the mu and kappa opioid receptors (Reisine and Pasternak, 1996). Under some training conditions these drugs were discriminated as similar to hydromorphone, whereas under other training conditions they were not discriminated as hydromorphone-like, as illustrated in Fig. 1. For example, in volunteers trained to discriminate i.m. hydromorphone versus saline (top panel), butorphanol, pentazocine, nalbuphine, and buprenorphine each fully substituted (i.e., produced $\geq 80\%$ of responding by pressing the drug-appropriate key) for the hydromorphone training condition (Preston et al., 1992). However, when tested in volunteers trained to discriminate among hydromorphone, saline, and a third choice of either pentazocine (middle panel) or butorphanol (lower panel), the agonist-agonists rarely substituted for hydromorphone (Preston et al., 1989; Preston and Bigelow, 1994). We have concluded from these studies that the outcome of human drug discrimination studies is dependent on the training conditions of the study and that the three-choice procedure is more useful for differentiating among drugs with overlapping discriminative stimuli than the two-choice, drug versus placebo, procedure.

A peculiarity in the research literature has been that, in humans, three-choice discrimination procedures have been needed to make the pharmacological differentiations among mu agonists and agonist-antagonists that animals make with two-choice procedures. One possible explanation for this difference is that the instructional and training procedures used with humans in two-choice discriminations have inadvertently established a different task from that trained in animals.

In the present study, we evaluated a different instructional set on the outcome of the two-choice discrimination procedure in humans. In the earlier two-choice study (Preston et al., 1992) participants were instructed to learn to discriminate between two different drugs labeled, for example, Drug

![Fig. 1](image-url).}
A and Drug B and to respond to test drugs by pressing the computer key of the most similar training drug. Participants were explicitly trained to discriminate hydromorphone and saline as Drug A and Drug B. In the instructional set of the present study, participants were instructed to learn to discriminate the drug labeled, for example, Drug A, and to respond on the “Drug A” key only if test drugs were identical with Drug A and to respond on a “Not Drug A” key for all other test drugs. They were then explicitly trained to discriminate hydromorphone and saline as Drug A and Not Drug A. Other than this instructional difference, the procedures and test conditions were the same for both studies.

Materials and Methods

Subjects. The participants were eight adult male volunteers with opioid-abuse histories but not physically dependent; they gave written informed consent and were paid for their participation. Ages ranged from 31 to 40 years (mean 37.8); weights ranged from 60.5 to 77.3 kg (mean 67.8). Subjects reported illicit opioid use of 6 to 24 years duration (mean 19.2). In addition to their opiate abuse, all reported current or past use of marijuana, alcohol, and cocaine; four of the eight subjects also reported current or past use of hypnotics. All reported current use of tobacco. None were physically dependent on opiates or alcohol at the time of the study as determined by self-report and observation while drug-free in the first 3 days of admission to the research unit. On the basis of physical examination, history, routine laboratory chemistries, and chest X-rays, participants were found to be in good health and without significant psychiatric disturbance other than their drug abuse. One additional participant enrolled but did not complete the study due to a medical condition that contraindicated the participation of the study drugs. His data are not included in this report.

Subjects participated while residing at a residential behavioral pharmacology research unit. The research unit contained a nursing station, patient bedrooms, recreational area, dining area, and experimental session rooms. Various recreational, reading, and craft materials and exercise equipment were available at all times other than during experimental sessions. Participants were free to smoke cigarettes except during sessions; no caffeinated beverages were available at the research unit.

Drugs. Drugs were administered i.m. under double blind conditions in a constant volume of 3 ml. The training drugs were saline (3 ml) and hydromorphone hydrochloride 3 mg/70 kg. Dose-response generalization training was conducted with hydromorphone (0, 0.125, 0.25, 0.5, 1, 2, 3, and 4 mg/70 kg), butorphanol (0, 0.375, 0.75, 1.5, 3, and 6 mg/70 kg), pentazocine (0, 4, 8, 16, 32, and 64 mg/70 kg), nalbuphine (0, 1.5, 3, 6, 12, and 24 mg/70 kg), and buprenorphine (0, 0.055, 0.11, 0.22, 0.45, and 0.9 mg/70 kg). Commercially available preparations of each drug were used: hydromorphone hydrochloride (10 mg/ml; Knoll Pharmaceuticals, Whippany, NJ); butorphanol tartrate (2 mg/ml; Bristol Laboratories, Syracuse, NY); pentazocine lactate (30 mg/ml; Winthrop Laboratories, Inc., New York, NY); nalbuphine hydrochloride (10 mg/ml; DuPont Pharmaceuticals, Wilmington, DE); buprenorphine hydrochloride (0.3 mg/ml; Norwich Eaton, Norwich, NY). Hydromorphone hydrochloride, butorphanol tartrate, and nalbuphine hydrochloride doses were calculated on the basis of the salts, and pentazocine lactate and buprenorphine hydrochloride doses were calculated as the drug base. Appropriate volumes of each drug were diluted to the desired concentration with bacteriostatic saline. For generalization testing a range of doses for each drug was selected such that the recommended analgesic dose fell mid-range.

General Methods. Subjects were informed: 1) that they could receive various psychoactive drugs that might have sedative properties, stimulant properties, opioid properties, or opioid blocking properties; 2) that the study involved evaluation of their ability to discriminate one drug from another and evaluation of the subjective, behavioral, and physiological effects of those drugs; and 3) that in each session they could earn money for correct discriminations. Subjects were trained to discriminate the presence of Drug A (hydromorphone 3 mg/70 kg) from the absence of Drug A. Subjects were instructed that they were to learn to identify a drug, identified to them as Drug A, from all others by the way it made them feel. They were instructed that they should respond on the Drug A key when the effects they experienced were exactly like those produced by Drug A and that they should respond on the Not Drug A key if the effects were not those of Drug A. Subjects were told that Drug A would not change during the study. Different letter codes were used for different subjects, but the A/not A terminology is used throughout for convenience.

A microcomputer presented all discrimination measures, questionnaires, and performance tests in a prearranged and timed sequence and printed and stored the data for each session. The subject indicated his responses on a numeric key pad.

The study proceeded in three phases, with sessions conducted once daily. Sessions were conducted in the same manner in all phases except for the information provided to the subject either before or after each session as described below. Discrimination training was conducted in sessions 1 through 4, during which the subject received, in randomized block order, two sessions of exposure to each of the two training conditions (hydromorphone 3 mg/70 kg and saline). During these training sessions before administration, hydromorphone was identified to the subject as Drug A and saline was identified to the subject as Not Drug A.

In sessions 5 through 8 acquisition of the discrimination was tested by exposing the subject to hydromorphone, 3 mg/70 kg, and saline twice in randomized block order. The purpose of these sessions was to determine whether the subject reliably identified hydromorphone as Drug A and saline as Not Drug A. During these and all subsequent exposures to the two training conditions, the subject received feedback about the correct response at the end of the session. This test of acquisition procedure was also interspersed among test sessions during the subsequent testing phase to provide continued training and to ensure continued correct discrimination.

Beginning with session 9, generalization test sessions were conducted. Test sessions were randomly interspersed with test of acquisition sessions, with approximately 52% of the sessions being test of acquisition and 48% test sessions. During this testing phase, dose-response curves for each active training drug were determined in randomized order, followed by dose-response curves for pentazocine, butorphanol, nalbuphine, and buprenorphine (in that order). Doses of active drug in each dose-response curve were administered in a randomized sequence. After each test session the subject did not receive feedback about the correct drug identification but was informed that it had been a test session and that the drug code could not be revealed.

Experimental Session. Daily sessions were conducted using methods similar to those previously reported (Preston et al., 1989). At the beginning of each experimental session, inspiration rate, pulse, temperature, blood pressure, and pupil diameter were recorded, and the subject completed baseline self-report questionnaires [adjective rating scales and the Addiction Research Center Inventory (ARCI) and the Digit Symbol Substitution Test (DSST) in the experimental room. The scheduled drug or saline was then injected. During the initial training sessions the subject was informed of the drug’s identifying letter code at the time of injection. The subject remained under staff observation for 20 min and then returned to the experimental room to complete the postdrug discrimination, subjective effect, and performance testing. Postdrug testing lasted for 90 min (20-110 min postinjection) as described in detail elsewhere (Preston et al., 1989). Subjective effects and DSST performance were assessed five times after drug administration; drug discrimination performance was assessed twice. At the end of the session the staff again recorded pupil diameter and other physiological measures. A
Hydromorphone/Not Hydromorphone Discrimination in Humans

sealed envelope was then opened, and the staff informed the patient of the letter-code identity of the administered drug or, on test sessions, that the code could not be revealed.

**Discrimination Procedures.** Drug discrimination data were collected in three ways. In each of these three procedures only correct responses were converted to monetary reinforcement for the subject. In Discrete Choice the subject indicated whether he thought he had received Drug A or Not Drug A as a single choice. In Point Distribution the subject distributed 50 points between the two training alternatives depending on how certain he was of the identity of the administered drug; this required typing in the numbers. On the Operant Responding measure the subject responded on a fixed interval, 1-s schedule on computer keys designated with Drug A or Not Drug A to earn points for 3 min; points (displayed on the computer screen) could be earned (at a maximum rate of one per second) for each of the training drugs by pressing the key corresponding to that drug.

The maximum amount of contingent payment available per session was approximately $10.00. Actual payment for correct responses was determined according to the following schedule: discrete choice measure, $3.00/session; point distribution measure, $0.03/point [100 points ($3.00/session)]; operant response measure, $0.011/point [approximately 360 points ($4.00/session)]. Subjects were not informed as to the precise monetary value of each response but were told that a bonus payment of up to $10.00 was available in each session and that the number of correct responses determined the bonus. Earnings for test sessions were calculated as the mean of the previous six tests of acquisition sessions. Earnings were reported at the end of each experimental session and paid to subjects at discharge from the study.

**Subject-Rated Measures.** Four questionnaires were completed: 1) visual analog scales (VAS), 2) a pharmacological class questionnaire, 3) an adjective rating scale, and 4) a shortened form of the ARCI. On the VAS, the subject ranked the degree of “Any Drug Effects,” “Liking,” “Good Effects,” “Bad Effects,” and “High” produced by the drug and rated the similarity of the drug to the training drugs on questions asking “How much is this drug like . . . (each of the training conditions, by letter code, e.g., Drug A, Not Drug A)” by placing an arrow along a 100-point line marked at either end with “none” and “extremely.” On the pharmacological class questionnaire, the subject categorized the drug effect as being most similar to one of 10 classes of psychoactive drugs (Preston et al., 1989). The adjective rating scale consisted of 32 items, which the subject rated on a 5-point scale from 0 (“no effect”) to 4 (“maximum effect”). The items in the adjective list were divided into three subscales (Preston et al., 1989): a 13-item Agonist scale, an 10-item Antagonist, and a 9-item Agonist-Antagonist opioid scale. The ratings of the individual items within each scale were summed to yield a single total score for each scale. The short form of the ARCI consisted of 49 true/false questions and contained five major subscales: Morphone-Benzodrine group (MBG, an index of euphoria); Pentobarbital, Chlorpromazine, Alcohol group (PCAG, an index of sedation); Lysergic Acid Diethylamine (LSD, an index of dysphoric and somatic changes); and Benzodrine group (BG) and Amphetamine scales (empirically derived amphetamine-sensitive scales) (Martin et al., 1971). At baseline, only the adjective rating scales and ARCI were completed; all questionnaires were completed at other time intervals as described above.

**Physiological and Performance Measures.** Physiological measures consisted of respiration rate, pulse, blood pressure, temperature, and pupil diameter. Vital signs were collected manually. Pupil diameter was measured from photographs taken in constant ambient room lighting using a Polaroid camera with 3× magnification. Psychomotor/cognitive performance was tested using a computerized version of the DSST developed in our laboratory (McLeod et al., 1982).

**Data Analysis.** Because time-related effects are minimal for these drugs within the 20- to 110-min postdrug assessment period, results of each session were summarized to yield a single value for each measure. Within-session means were calculated for VAS. Percent Drug A- and Not Drug A-appropriate responses were calculated for each discrimination measure. Mean changes from predrug scores were calculated for the ARCI, adjective rating scales, and DSST scores and physiological measures.

Within-session means (or mean changes from baseline) from two exposures to hydromorphone and saline (sessions 5–8) were analyzed using two-factor, repeated measures ANOVA (with factors of training drug and session) to evaluate the effects of the training drugs in the test of acquisition phase of the study. Session means were used for ease of data presentation and because time-related effects were small.

Within-session means (or mean changes from baseline) from the one exposure to each generalization test condition were used to determine dose responses from the generalization testing. Dose-response functions for each test drug (including its appropriate saline control session) were analyzed separately using one-factor, repeated measures analyses of variance to test the main effect of dose. An overall analysis comparing all five drugs together used a two-factor, repeated measures analysis of variance (with factors of drug and dose). To have equal numbers of doses across drugs, the 0.125 and 3 mg/70 kg doses of hydromorphone were excluded. Post hoc between-drug comparisons of the highest dose(s) to hydromorphone, 4 mg/70 kg, were made using the Tukey honestly significant difference (HSD) test.

The method of Finney (1964) was used to assess whether the dose effect curves of each of the four agonist-antagonist test drugs satisfied the requirements of linearity and parallelism to hydromorphone necessary for calculating relative potency. These calculations were made for 11 variables—seven VAS scales (Drug Effect, Liking, Good Effects, Bad Effects, High, Like Drug A, and Not Like Drug A), two drug class identification responses (identifications as opiate, identifications as placebo), and two operant drug discrimination measures (responses as Drug A, responses as Not-Drug A). The calculated values are expressed relative to morphine, because it is the most commonly used standard among opioids; 1.3 mg of hydromorphone was considered equivalent to 10 mg of morphine (Reisine and Pastersak, 1996).

Conservative *F* tests employing Huynh-Feldt probability levels were used to interpret the results of all analyses of variance. Effects were considered statistically significant if *P* was <.05. *P* values greater than .05 and less than or equal to .1 are presented to indicate trends under *Results*. ANOVA were conducted using SPSS statistical software (SPSS Inc., Chicago, IL).

**Results**

**Test of Acquisition.** Subjects reliably discriminated hydromorphone as Drug A and saline as Not Drug A in the test of acquisition sessions after training (sessions 5–8) with 85 to 94% correct responding for each substance. Two-way ANOVA showed that there was a significant main effect of drug but no main session effect or drug by session interaction (Table 1). The results were similar on all three discrimination measures.

The training dose of hydromorphone (3 mg/70 kg) produced characteristic changes on subjective, physiological, and psychomotor performance measures in these sessions (Table 1). Ratings of hydromorphone on the VAS “How much is this drug like Drug A (hydromorphone)?” was significantly higher than ratings of saline. The reverse was true on the “How much is this drug like Not Drug A” scale, with saline producing higher ratings than hydromorphone. Of the subjective effect measures, significant main effects of drug were produced on the “Any Drug Effect,” “High,” “Liking,” and “Good Effects” VAS. On the ARCI, hydromorphone produced signif-
icant increases in the Amphetamine scale and a trend (P < .10) toward increases in the MBG and BG scales. Hydromorphone produced significant increases in the Agonist and Agonist/Antagonist adjective rating scales compared with saline, although its effect on the Agonist/Antagonist scale was quite small. A number of individual items (itching, turning of stomach, nodding, relaxed, dry mouth, and tingling) also showed significant main effects of drug, with higher ratings after hydromorphone administration. On physiological measures, hydromorphone significantly decreased pupil diameter. There were no significant differences between hydromorphone and saline on other physiological measures or DSST performance.

**Generalization Testing.** Results of the generalization testing are summarized in Tables 2 and 3, and selected variables are shown in Figs. 2 through 4. Table 2 summarizes significant effects as well as major trends for P values less than or equal to .10. The arrows indicate the directions of drug effects relative to placebo.

As in the training and test of acquisition sessions, similar results were found on all three discrimination measures (Table 2); results of the operant response discrimination measure are shown in Fig. 2. Criterion for full substitution for the training conditions was set at 80% (mean across all participants) drug-appropriate responding. Because most participants responded in an all-or-none fashion, the 80% criteria corresponds to seven of eight participants discriminating the test drug as hydromorphone-like. Hydromorphone doses of 3.0 and 4.0 mg were identified as hydromorphone in 80% or more of trials. Pentazocine, 32 mg, produced 80% or greater identifications as hydromorphone, although hydromorphone-appropriate responses were decreased at 64 mg, resulting in an inverted U-shaped dose-response curve. Neither butorphanol nor nalbuphine produced greater than 80% hydromorphone-appropriate responses at any test dose; there were no statistically significant effects on either hydromorphone- or not hydromorphone-appropriate responses with either drug. Buprenorphine produced dose-related increases in hydromorphone-appropriate responses, with 94% at 0.9 mg.

The dose-response curves of the five test drugs on the “How much is this drug like Drug A” (hydromorphone) VAS (Fig. 2) resembled those of the behavioral discrimination results,
although less so for butorphanol and nalbuphine. All but butorphanol produced significant increases in this VAS rating (Table 2). Hydromorphone produced dose-related increases to 100% similarity to the hydromorphone training condition. All four agonist-antagonists produced maximal scores less than 100%; pentazocine produced an inverted U-shaped curve, and maximal scores produced by butorphanol, nalbuphine and buprenorphine were in the range of 30 to 60% of the total scale. All five drugs decreased ratings on the “How much is this drug like Not Drug A” VAS (Fig. 2), although the decrease with nalbuphine did not reach statistical significance (P < .08).

On the pharmacological class questionnaire (Table 3) sali- line and the lowest doses of each of the test drugs were primarily identified as placebo: 88 to 100% of occasions for saline; 100% for hydromorphone (0.125 mg); 94% for pentazocine (4 mg); 78% for butorphanol (0.375 mg); 75% for nalbuphine (1.5 mg); 84% for buprenorphine (0.055 mg). Hydromorphone produced dose-related increases in identifications as an opioid, with 4 mg being identified as an opioid in 100% of opportunities. Pentazocine produced an inverted U-shaped dose-related curve in identifications as an opioid (to a maximum of 69% at 32 mg) and a decrease in identifications as an opioid (38%) at 64 mg. Butorphanol produced some identifications as an opioid (maximum 56% at 3.0 mg) but was identified as being similar to a variety of classes (primarily benzodiazepines or barbiturate). Nalbuphine was identified primarily as opiate-like (maximum 62% at 12 and 24 mg) but also as benzodiazepine/barbiturate-like at higher doses (25% at 12 and 24 mg). Buprenorphine produced dose-related increases in identifications as an opioid with the highest dose tested (0.9 mg) being identified as an opioid in 91% of oppor-

The effects of the five test drugs on all of the VAS are summarized in Table 2; graphs of selected measures are shown in Fig. 3. All five test drugs significantly increased ratings on the “Any Drug Effect” and “High” VAS. Differential effects of the drugs were found on the “Liking,” “Good Effects,” and “Bad Effects” VAS. Hydromorphone, pentazocine, and buprenorphine each significantly increased “Lik-
produced significant increases in ratings on the Agonist scale and the individual items (skin itching, nodding, and dry mouth (from the Agonist scale) and tingling (from the Agonist-Antagonist scale)).

With a few exceptions, the five test drugs produced similar effects on the physiological and psychomotor performance measures (Table 2). All five test drugs significantly decreased pupil diameter (Fig. 4); none of the five test drugs had significant effects on temperature, blood pressure, or respiration rate. Buprenorphine significantly increased heart rate; nalbuphine tended to decrease heart rate, although this did not reach statistical significance. Butorphanol significantly decreased the number of correct trials and number of trials completed on the DSST (Fig. 4). Buprenorphine also tended to decrease number of correct trials, although this did not reach statistical significance. Hydromorphone, pentazocine, and nalbuphine had no significant effects on performance on the DSST.

Across-drug ANOVA and Tukey HSD post hoc tests were conducted to provide more direct between-drug comparisons of the qualitative and quantitative effects at the doses used to determine whether the test drugs substituted for the hydromorphone training condition (Table 2). In post hoc testing, butorphanol, 6 mg, differed from hydromorphone, 4 mg, on a number of measures, including significantly lower hydromorphone-appropriate responding and significantly greater not hydromorphone-appropriate responding on all three discrimination measures. This is consistent with the observed significant qualitative differences in subjective effects, including lower ratings on measures of like hydromorphone, liking, good effects, and MBG scale and higher scores on the antagonist and agonist-antagonist adjective rating scales. In addition to the measures shown in Table 2, butorphanol differed significantly from hydromorphone on 11 individual adjective items, including higher ratings of feeling sleepy, drunken, shaky, tired, restless, confused, and lightheaded and lower ratings of itching, talkative, drive, and energetic. Conversely, buprenorphine, 0.9 mg, differed on two measures, the like hydromorphone VAS and ratings of feeling itchy, both rated lower than hydromorphone, 4 mg. Nalbuphine, 24 mg, also produced relatively few significant differences from hydromorphone, 4 mg, including lower scores on five subjective effects measures listed in Table 2 and one item on the adjective rating scale, itchy, indicating more quantitative than qualitative differences from hydromorphone. Pentazocine, 32 mg, did not significantly differ from hydromorphone, 4 mg, on any of the VAS but did produce significantly lower ratings on the antagonist adjective rating scale and two individual items, itching and dry mouth; in contrast, pentazocine, 64 mg, had a greater number of differences, including significantly increased ratings of shaky and lower ratings of talkative, itchy, and dry mouth, consistent with less similarity to the 4-mg dose of hydromorphone than the 32-mg dose.

In the relative potency analyses of pentazocine none of the 11 variables satisfied the criteria—failing in each case to be linear. For butorphanol, nalbuphine, and buprenorphine, 5, 6, and 6 variables, respectively, satisfied the criteria. For those measures the mean estimates for equivalence to 10 mg parenteral morphine were butorphanol (2.17 mg), nalbuphine (10.00 mg), and buprenorphine (0.40 mg). These calcu-

### TABLE 3

**Results of the pharmacological class identification questionnaire**

Percentage of identifications of each test drug dose out of 32 responses (four responses for each of eight subjects). Identifications as phenothiazines, stimulants, phencyclidine, and hallucinogens were combined and listed as Other. Not all rows add to 100% due to rounding.

<table>
<thead>
<tr>
<th>Test Drug</th>
<th>Placebo</th>
<th>Opioid</th>
<th>Barbiturate/Benzodiazepine</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hydromorphone</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.125</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>84</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>63</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>38</td>
<td>59</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>38</td>
<td>63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pentazocine</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>94</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>97</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>59</td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>25</td>
<td>69</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>31</td>
<td>38</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td><strong>Butorphanol</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.375</td>
<td>78</td>
<td>9</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>0.75</td>
<td>72</td>
<td>16</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>28</td>
<td>38</td>
<td>28</td>
<td>6</td>
</tr>
<tr>
<td>3.0</td>
<td>13</td>
<td>56</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>6.0</td>
<td>13</td>
<td>41</td>
<td>34</td>
<td>13</td>
</tr>
<tr>
<td><strong>Nalbuphine</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>88</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>75</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>53</td>
<td>47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>38</td>
<td>38</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>13</td>
<td>63</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td><strong>Buprenorphine</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.055</td>
<td>84</td>
<td>3</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>0.11</td>
<td>97</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.22</td>
<td>66</td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.45</td>
<td>19</td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td>9</td>
<td>91</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ing” and “Good Effects” scales, with no increases in ratings of “Bad Effects.” Butorphanol tended ($P < .10$) to increase the “Good Effects” and “Bad Effects” scales, with no effect on the “Liking” scale. Nalbuphine had no significant effects on any of the three scales.

The five test drugs were also differentiated on the ARCI scales and adjective rating scales (Table 2). Hydromorphone produced significant increases on the MBG and Amphetamine scale of the ARCI, the Agonist adjective rating scale, and on six individual adjective items: itchy, relaxed, nodding, talkative, and dry mouth (from the Agonist scale), and tingling (from the Agonist-Antagonist scale). Pentazocine significantly increased the Amphetamine scales of the ARCI and tended to increase the Agonist adjective rating scale but had no significant effects on any of the individual items or on the ARCI subscales. Butorphanol had no significant effects on ARCI, although it did significantly increase ratings on the Agonist adjective rating scale scores and one individual item in the adjective rating questionnaire: nodding (from the Agonist scale). Nalbuphine significantly increased ratings on the Amphetamine ARCI scale, the Agonist adjective rating scales, and on two individual items in the adjective rating questionnaire: nodding (from the Agonist scale) and tingling (from the Agonist-Antagonist scale). Buprenorphine significantly increased the BG scale of the ARCI; it also tended to increase scores on the MBG and Amphetamine scales but did not reach statistical significance. On the adjective rating questionnaire buprenorphine
lated values are in very close agreement with the published relative potency values of 2, 10, and 0.4, respectively (Reisine and Pasternak, 1996).

**Discussion**

In the present study, subjects were trained to discriminate Drug A (hydromorphone 3 mg/70 kg administered i.m.) from Not Drug A (saline) and tested with a series of opioid agonist-antagonists. Only hydromorphone and buprenorphine produced significant dose-related increases in hydromorphone-appropriate responses. Pentazocine also produced significant increases, with greater than 80% at 32 mg, but the highest dose tested (64 mg) produced fewer hydromorphone-appropriate responses, resulting in an inverted U-shaped dose-response curve. Neither butorphanol nor nalbuphine produced dose-related or statistically significant increases in hydromorphone-appropriate responses. Thus, in this procedure in which subjects were trained to discriminate hydromorphone as Drug A and instructed to identify all other drugs as Not Drug A and trained to respond on the Not Drug A key with saline, there was clear differentiation among the

---

**Fig. 2.** Dose-response curves of percentage hydromorphone-appropriate responding on the operant response discrimination measure and ratings on the “How much is this drug like Drug A (hydromorphone)?” and the “How much is this drug like Not Drug A?” VAS produced by hydromorphone, pentazocine, butorphanol, nalbuphine, and buprenorphine. Data are the mean of two assessments (at 60 and 80 min postdrug administration) in each session. Each point is the mean ± 1 S.E. based on one administration of each drug condition in each of eight subjects. The S.E.M. is not shown where it is less than the radius of the symbol.

**Fig. 3.** Dose-response curves of hydromorphone, pentazocine, butorphanol, nalbuphine, and buprenorphine on the Drug Effect, Liking, Good Effects, and Bad Effects VAS. Data are the mean of five assessments in each session. Each point is the mean ± 1 S.E. based on one administration of each drug condition in each of eight subjects. The S.E.M. is not shown where it is less than the radius of the symbol.
four agonist-antagonists on the behavioral discrimination measures.

In a previous two-choice study in which subjects were explicitly trained with hydromorphone as Drug A and saline as Drug B (for example) with instructions to respond on the key of the most similar drug, all five test drugs fully substituted for the hydromorphone training condition (Fig. 1, top panel; Preston et al., 1992). Those results are quite different from the present study as well as from several three-choice studies conducted in our laboratory. For example, none of the studied agonist-antagonists fully substituted for hydromorphone in volunteers trained to discriminate among hydromorphone, saline, and pentazocine (Fig. 1, middle panel; Preston et al., 1989), and only buprenorphine fully substituted in volunteers trained to discriminate among hydromorphone, saline, and butorphanol (Fig. 1, bottom panel; Preston and Bigelow, 1994). Furthermore, in subjects trained to discriminate among hydromorphone at 0, 1, and 4 mg, pentazocine, butorphanol, and nalbuphine fully substituted only for the hydromorphone 1-mg training dose, whereas buprenorphine produced intermediate responding and fully substituted for neither 1 nor 4 mg (Jones et al., 1999). Based on the results of these prior studies, we have concluded that the outcome of human drug discrimination studies is dependent on the discrimination training conditions (including number and choice of training drugs used) (Jones et al., 1999). The present study demonstrates that instructions are one of these important procedural factors that must be considered in the design, conduct, and interpretation of drug discrimination studies in humans. Although based on across-study comparisons, the conclusion is strengthened by the fact that the series of studies was conducted in a single laboratory with similar methods and subjects across studies. Research by Smith and Bickel (1999) showing that an instruction-based, novel-response procedure has increased pharmacological selectivity compared with a standard two-choice procedure in human sedative discrimination also supports this conclusion.

Instructions are a feasible experimental manipulation only in human studies. Nevertheless, animal studies also support the more general finding that a number of procedural factors, such as training dose, species, training drug, and context affect discrimination performance, including the differentiation of mu and kappa opioids (Woods et al., 1988; Jarbe, 1989; Ator, 1999). The fact that discrimination performance is influenced by instructions offers the tantalizing possibility that the degree of correspondence between results of animal versus human studies might shed light on what concept or construct is discriminated by animals. This has long been a topic of debate (Overton et al., 1983; Jarbe, 1989).

Given that different generalization profiles are found with different discrimination procedures, the question arises as to which procedure best reflects the receptor activities of the test drugs. One comparison that might be made is to assays in vitro. Mu and kappa receptors are coupled to the G-protein second messenger system; binding of an agonist to the receptor activates the G-protein by stimulating the release of GDP and the binding of GTP. All four agonist-antagonists have at least partial agonist activity at both receptor types and have been tested for their ability to stimulate [35S]GTP binding in cloned mu and kappa receptors (Emmerson et al., 1996; Zhu et al., 1997; Selley et al., 1998; Remmers et al., 1999). Butorphanol, nalbuphine, and pentazocine were reported to have similar low efficacies in a C6 glioma cell line expressing the mu receptors (Emmerson et al., 1996). Selley et al. (1998) determined the intrinsic activity of a series of opioids in mu receptor-transfected Chinese hamster ovary (CHO) cells and rat thalamus; they found buprenorphine to be a partial agonist with moderate efficacy (like meperidine) and nalbuphine low efficacy (like nalorphine and levallorphan). In an assay of kappa agonist activ-
ity in CHO cells, Zhu et al. (1997) found a rank order of potencies of pentazocine = nalbuphine > buprenorphine in stimulating \( ^{35}S \text{GTP} \gamma S \) binding; pentazocine was a partial agonist with maximal responses 23% of full agonists, and nalbuphine and buprenorphine had low levels of agonist activities. Butorphanol had 22% efficacy and nalbuphine had 18% (maximal stimulation divided by maximal stimulation of a full agonist) in a C6 glioma cell line expressing kappa receptors (Remmers et al., 1999). Overall, it appears that buprenorphine has greater mu agonist activity and lower kappa agonist activity than the other three. The relative activities of pentazocine, butorphanol, and nalbuphine at mu and kappa receptors is less clear, although all appear to have low to moderate activity at both receptors.

The results of the present study are generally consistent with the intrinsic activities of the agonist-antagonists tested at the mu and kappa opioid receptors in vitro, with buprenorphine fully substituting for the mu agonist, whereas pentazocine did so at one dose and butorphanol and nalbuphine not at all (Fig. 2). In contrast, all four agonist-antagonists fully substituted for hydromorphone in the two-choice hydromorphone-saline (Drug A versus Drug B) discrimination, which differed from the present study only by instructions (top panel, Fig. 1; Preston et al., 1992). Thus, the hydromorphone/not hydromorphone discrimination resulted in greater pharmacological specificity across opioid receptor types than the hydromorphone-saline discrimination.

Human studies permit assessment of the relationship between discrimination responding and subjective effect measures. The present hydromorphone/not hydromorphone discrimination results showed good agreement with subjective effect measures, particularly as assessed by the Tukey HSD analyses. Butorphanol, which did not substitute for hydromorphone on the discrimination measures, was differentiated from hydromorphone on a substantial number of subjective effect measures, including some effects consistent with kappa agonist activity. Buprenorphine, the only test drug to fully substitute, had few subjective effects significantly different from hydromorphone, and none of them suggested qualitative differences. The inverted U-shaped dose-response curve produced by pentazocine on the discrimination measures was mirrored by changes in subjective effect measures. The subjective effects shown in the present study are consistent with earlier studies showing that mu agonists and mixed action opioids (drugs with mu and kappa agonist activity) produce different patterns both of discrimination and subjective effects (Dykstra et al., 1997). The present hydromorphone/not hydromorphone discrimination results agree better with subjective effect measures than did the hydromorphone-saline discrimination results (Preston et al., 1992) and are similar to results seen with our earlier three-choice studies (Preston et al., 1989; Preston and Bigelow, 1994). As noted above, the hydromorphone-saline discrimination did not differentiate among the agonist-antagonists (Preston et al., 1992; Fig. 1), with each of agonist-antagonists substituting for hydromorphone, even though, as in the present study, the subjective effects profiles varied across the five study drugs. Interestingly, subjects were less likely to identify butorphanol and nalbuphine as opiate-like in the hydromorphone/not hydromorphone discrimination study than in the hydromorphone-saline discrimination study, suggesting that the discrimination categories used by participants influenced their identifications on the pharmacological class questionnaire. Thus, there is good agreement between discrimination responding and subjective effects, particularly with the Drug A/Not Drug A and three-choice discrimination procedures, although there may be some reciprocal influences between discrimination training and subjective effect measures.

Both types of measures, drug discrimination and subjective effects, have utility for evaluating the stimulus effects of drugs in humans. Drug discrimination provides a good mechanism for comparison across drugs, whereas subjective effect measures provide descriptive information about the stimuli individuals experience after drug administration. Determination of the similarity of two drugs based only on subjective effects measures can be difficult because of the number of individual measures that are collected and the unknown contribution of each to the overall effect. On the other hand, subjective effect measures are useful in the interpretation of discrimination results, because drug discrimination responding is significantly influenced by the training conditions. In addition, studies collecting discrimination and subjective effects of drugs in humans have provided evidence supporting the utility of animal drug discrimination paradigm as a model for subjective effects in humans.

In conclusion, the agonist-antagonists were differentially discriminated as hydromorphone-like, with buprenorphine fully substituting, pentazocine fully substituting at one dose, but not at a higher dose, and butorphanol and nalbuphine not fully substituting for hydromorphone in this two-choice, hydromorphone/not hydromorphone discrimination procedure. The results differ from those of a previous two-choice, hydromorphone-saline discrimination study, confirming that instructions to subjects are an important factor in human drug discrimination studies. Thus, the present study confirmed the significant impact of training conditions on the outcome of drug discrimination studies and demonstrated that the two-choice drug/not drug instructional set produced discrimination results that are more consistent with the intrinsic activities of these drugs in in vitro assays of mu and kappa agonist activity and more consistent with their subjective effects than does the drug A/drug B procedure.

Acknowledgments

We thank Rebecca Fromme for technical assistance, Tim Mudric and Linda Felch for statistical analysis, John Yingling for computer programming and equipment set-up, the residential nursing staff for assistance with volunteers and sessions, and the recruiting and assessment staff for assistance with screening and enrollment of volunteers.

References


