

Effort-Based Decision-Making Paradigms for Clinical Trials in Schizophrenia: Part 2—External Validity and Correlates

William P. Horan^{*1,2}, L. Felice Reddy^{1,2}, Deanna M. Barch³, Robert W. Buchanan⁴, Eduardo Dunayevich⁵, James M. Gold⁴, Steven R. Marder^{1,2}, Jonathan K. Wynn^{1,2}, Jared W. Young^{6,7}, and Michael F. Green^{1,2}

¹Department of Veterans Affairs VISN 22 Mental Illness Research, Education, and Clinical Center, Los Angeles, CA; ²UCLA Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine, Los Angeles, CA; ³Departments of Psychology, Psychiatry and Radiology, Washington University, St. Louis, MO; ⁴Maryland Psychiatric Research Center, Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD; ⁵Amgen Inc., Thousand Oaks, CA; ⁶Department of Psychiatry, University of California San Diego, San Deigo, CA; ⁷Research Service, VA San Diego Healthcare System, San Diego, CA

* To whom correspondence should be addressed; William P. Horan, Ph.D. VA Greater Los Angeles Healthcare System, MIRECC 210A, Building 210, 11301 Wilshire Boulevard, Los Angeles, CA 90073, US; tel: 310-478-3711, fax: 310-268-4056, e-mail: horan@ucla.edu

Effort-based decision making has strong conceptual links to the motivational disturbances that define a key subdomain of negative symptoms. However, the extent to which effort-based decision-making performance relates to negative symptoms, and other clinical and functionally important variables has yet to be systematically investigated. In 94 clinically stable outpatients with schizophrenia, we examined the external validity of 5 effort-based paradigms, including the Effort Expenditure for Rewards, Balloon Effort, Grip Strength Effort, Deck Choice Effort, and Perceptual Effort tasks. These tasks covered 3 types of effort: physical, cognitive, and perceptual. Correlations between effort related performance and 6 classes of variables were examined, including: (1) negative symptoms, (2) clinically rated motivation and community role functioning, (3) self-reported motivational traits, (4) neurocognition, (5) other psychiatric symptoms and clinical/demographic characteristics, and (6) subjective valuation of monetary rewards. Effort paradigms showed small to medium relationships to clinical ratings of negative symptoms, motivation, and functioning, with the pattern more consistent for some measures than others. They also showed small to medium relations with neurocognitive functioning, but were generally unrelated to other psychiatric symptoms, self-reported traits, antipsychotic medications, side effects, and subjective valuation of money. There were relatively strong interrelationships among the effort measures. In conjunction with findings from a companion psychometric article, all the paradigms warrant further consideration and development, and 2 show the strongest potential for clinical trial use at this juncture.

Key words: schizophrenia/effort/decision making/functional outcome

Introduction

Negative symptoms are treatment-refractory features of schizophrenia that have long been associated with poor functional outcome. Clinical trials of novel drugs intended to improve negative symptoms rely primarily, and often exclusively, on clinical interview-based outcome measures. Interview-based measures, however, are prone to influences that can complicate assessment, such as recall difficulties/biases, unwillingness to disclose, poor verbal skills, and lack of insight. Treatment research stands to benefit from behavioral paradigms, which could provide more objective and sensitive measures of motivational disturbances. A particularly rich foundation for such translational research is the extensive preclinical literature on effort-based decision making, which operationalizes motivation as the amount of effort an animal is willing to exert for different levels of reward (Young J. W. and Markou A.,¹ this issue).

The current report is the fourth in a set of articles that describes multidisciplinary translational research to evaluate the suitability of 5 effort-based decision-making paradigms for use in clinical trials. A companion article (Reddy L. R. et al.,² this issue) describes the paradigms and evaluates patient vs control differences, test-retest reliability, utility as repeated measures, and tolerability. Because new interventions are intended to ultimately improve clinical and/or functional outcomes, it is important to examine the external validity of these tasks, particularly their relations to clinically meaningful measures of motivation and functioning.

Eight published studies have evaluated effort-based decision-making tasks and their correlates in

schizophrenia.³ Relations with negative symptoms have been examined in all of these studies and the results have been inconsistent, with some,^{4,8} though not all,⁹⁻¹¹ finding significant associations. Notably, various analytic approaches have been used to examine negative symptoms, including treating them as continuous total or subscale (eg, experiential negative symptoms) ratings, and subgroup approaches (median splits). Findings have been similarly equivocal regarding correlations with community functioning/quality of life, neurocognition, other psychiatric symptoms, and self-reported motivational/emotional traits. Additionally, it is difficult to integrate findings across studies due to the considerable variability in methods and sample sizes, and because no study included more than a single effort paradigm.

Other factors could confound links between behavioral and clinical measures of negative symptoms, including the type of antipsychotic medication and potential group differences in the subjective valuation of money (ie, in how much monetary rewards are valued^{9,10}). To date, no associations have been found across studies for type/dose of medication or subjective valuation, although only 2 studies have considered the latter.

The goal of this article is to characterize the relationship between 5 effort-based paradigms and clinically and functionally relevant variables in the largest sample of individuals with schizophrenia studied to date. We conducted analyses on 4 sets of variables to evaluate: (1) Convergent validity: the primary indexes were clinically rated experiential negative symptoms and intrinsic motivation; (2) Associations with real world functioning; (3) Associations with self-reported traits related to motivation; (4) Associations with potential confounds that could influence interpretation of results, including neurocognition, other clinical/demographic characteristics, antipsychotic medication type and side effects, and subjective valuation of money. We also explored interrelationships among the 5 effort paradigms.

Methods

Participants

Participant characteristics, recruitment, inclusion/exclusion criteria, and psychometric properties of the effort paradigms are described in a companion article (Reddy L. R. et al,² this issue). Briefly, participants included 94 outpatients with schizophrenia and 40 demographically matched healthy controls. Diagnosis was determined by the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition-I and -II.^{12,13} Effort paradigms were administered twice (baseline and 4-week retest) with order counter-balanced across participants. The 5 paradigms took approximately 2½ h to complete and were well tolerated. Only data from the baseline assessment were used in the current analyses. With the exception of

1 task (subjective valuation of money index, described below), data from healthy controls are not presented here. After providing a complete description of the study to prospective study participants, written informed consent was obtained prior to participation.

Effort-Based Decision-Making Tasks

The 5 paradigms and their theoretical background are fully described in a companion article (Reddy L. R. et al,² this issue). The computerized paradigms included 1 for cognitive effort, 1 for perceptual effort, and 3 for physical effort, which all involve making a series of choices between hard vs easy tasks that involve varying levels of monetary reward contingencies. The cognitive task was the Deck Choice Effort Task, which assesses willingness to perform hard judgments (alternating between judgments of odd/even and $>/< 5$ on each trial) vs easy judgments (making the same judgment of either odd/even or $>/< 5$ across all trials) for sets of serially presented numbers across 3 levels of reward. The perceptual task was the Perceptual Effort Task, which involves identifying the location of a briefly presented target visual stimulus for high-effort tasks (low contrast between target and background) vs low-effort tasks (high contrast between target and background) across 4 levels of reward.

The physical effort tasks included: Grip Effort Task, which assesses willingness to make hard hand grips (90% of personal maximum) vs easy hand grips (50% of personal maximum) across 3 levels of reward; Balloon Effort Task,⁵ assessing willingness to perform hard (100 button presses) vs easy (10 button presses) tasks to inflate and pop a graphically presented balloon across 5 reward levels; and Effort Expenditure for Rewards Task (EEfRT),¹⁴ which assesses willingness to perform hard (larger number of button presses in a relative short time) vs easy (small number of button presses in a relatively long time) tasks, this task incorporates 3 different reward levels, as well as 2 probability levels.

After examining a variety of potential indices, we selected a difference score as the representative index for each paradigm (based on psychometric properties; Reddy L. R. et al,² this issue), defined as % hard choices during the highest reward condition – % hard choices during the lowest reward condition. Higher scores indicate greater willingness to exert effort for large vs small rewards. Because the EEfRT task also manipulates probability level, we calculated a difference score for probability levels (“EEfRT Probability”), which equals % hard choices in the high probability condition – % hard choices in the low probability condition.

Difference scores create a problem for “inflexible responders” who selected either 100% hard tasks across all reward levels or 100% easy tasks across all reward levels on each of the effort paradigms (Reddy L. R. et al,² this issue). With difference scores, participants who

select either all hard or all easy would both be assigned a value of “0” (ie, no difference between the highest and the lowest reward levels). Because these 2 subgroups of participants reflect qualitatively different response profiles and willingness to exert effort for rewards, it is a problem to give these very different subtypes the same value. For this reason we removed from analyses participants who made only hard selections across all reward levels. This removed from analysis subjects who had no room to demonstrate increases in effort allocation. Removing these inflexible responders resulted in the following sample sizes: Deck Choice Effort: 78; Perceptual Effort: 82; Balloon Effort: 69; Grip Effort: 84; and EEfRT: 85.

Negative Symptoms

Given the current lack of consensus about the optimal assessment method for negative symptoms,³ multiple interview-based assessments were included. Participants received the Clinical Assessment Interview for Negative Symptoms (CAINS),¹⁵ a recently developed scale that includes 2 subscales. The Motivation and Pleasure (MAP) subscale includes 9 items based on motivation, interest, and emotional experiences, as well as reported engagement in relevant social, vocational, and recreational activities, over the past week. The Expression (EXP) subscale includes 4 items based on interviewer ratings of affective and verbal expression. Based on the conceptual link between effort-based decision making and the experiential negative symptoms, the CAINS MAP was our primary index.

Two additional measures that have been more widely used in clinical trials were also administered. The Negative Symptom Assessment (NSA-16)¹⁶ contains 16 items that cover 5 factors: communication, emotion/affect, social activity, motivation, and psychomotor activity. The global ratings score was used. The Positive and Negative Syndrome Scale (PANSS)¹⁷ Negative Symptom factor score, which focuses on expressive symptoms, was also used.

Motivation

We supplemented our clinical assessment of motivation with 3 items from the Intrapsychic Foundations subscale of the Quality of Life Scale¹⁸: Sense of Purpose, Motivation, and Curiosity, which have been described as an index of Intrinsic Motivation.¹⁹⁻²² A mean score was used, with higher scores indicating higher levels of motivation.

Community Role Functioning

To assess current functioning (past month), we administered the Role Functioning Scale (RFS),²³ which includes separate ratings for different domains of functioning. Ratings were based on a semi-structured interview with standardized probe questions. We included ratings for Working Productivity, Independent Living, and Social/

Family Relationships. Higher RFS scores indicate better functioning.

Self-Reported Traits

Participants completed 2 measures of traits that are conceptually related to negative symptoms and effort-based decision making. First, the behavioral inhibition system/behavioral approach system (BIS/BAS) scale²⁴ is a 24-item self-reported measure with 2 subscales that assess the sensitivity of 2 general motivational systems: (1) the BIS scale measures a behavioral avoidance (or inhibition) system that regulates aversive motivation in which the goal is to avoid unpleasant stimuli; (2) the BAS scale (comprised of 3 subscales) measures a behavioral approach system that regulates appetitive motives in which the goal is to move toward desired stimuli. Higher scores indicate greater BIS and BAS sensitivity. Second, the Defeatist Performance Attitude Scale (DPAS)²⁵ is a 15-item self-reported scale that measures defeatist attitudes about one's ability to perform goal-directed tasks. Higher total scores indicate more severe defeatist performance attitudes.

Neurocognition

Neurocognition was assessed using the MATRICS Consensus Cognitive Battery (MCCB).²⁶ Although we expected that neurocognition might correlate with performance on the effort tasks (see^{5,27-30} for discussions of the relationship between neurocognition and negative symptoms) we would be concerned if the correlations were very high (eg, >.60) as this would suggest the effort tasks were not assessing a meaningfully different construct. The MCCB includes 10 tests to measure 7 domains of cognition, including: speed of processing, attention/vigilance, working memory, verbal memory, visual memory, reasoning and problem solving, and social cognition. Standardized *T*-scores were computed for each domain, correcting for age and gender. The composite score was based on the average *T*-score from each of the domains.

Other Psychiatric Symptoms

Patients were administered the PANSS.¹⁷ In addition to negative symptoms, we examined the positive, depression/anxiety, and disorganization subscales. Clinical interviewers were trained on all the symptom assessment instruments through the Treatment Unit of the Department of Veterans Affairs VISN 22 Mental Illness Research, Education, and Clinical Center. Symptom raters were trained to a minimum intra-class correlation coefficient (ICC) of .80 on relevant subscales.

Side Effect Ratings

Since antipsychotic-induced motor disturbances are a potential confound for effort-based tasks, participants were administered the Brief Clinical Assessment of

Movement Disorders,³¹ a structured extrapyramidal side effects assessment of akathisia, rigidity, and tardive dyskinesia. Raters received training from an experienced research psychiatrist (SRM). Each symptom was rated from 0 (none) to 3 (severe) and an overall score was calculated.

Subjective Value of Money Index

For subjective valuation of money, participants were asked to: “rate how valuable (ie, how important) the following amounts of money are to you” (based on^{32,33}). Participants rated 7 monetary amounts (US\$ 10, 20, 50, 100, 200, 500, 1000) on a scale from 0 (not at all valuable) to 10 (extremely valuable). The rating for \$10 was subtracted from the rating for \$1000 to represent subjective sensitivity to gradations in monetary value; the lower the value, the less the sensitivity (ie, more similar value ratings from highest and lowest amounts).

Statistical Analyses

First, descriptive statistics were computed for each of the variables. Second, 4 sets of analyses were conducted to evaluate: (1) Convergent validity with the CAINS MAP and QLS Intrinsic Motivation scales; (2) Correlations with RFS real world functioning scales; (3) Correlations with self-reported traits; and (4) Associations with potential confounds that could influence interpretation of results, including neurocognition, other clinical/demographic characteristics, antipsychotic medication type and side effects, and subjective valuation of money. Because this was a psychometric and validation study, rather than a hypothesis testing study, corrections for multiple statistical tests were not used.³⁴ Because several of the variables had non-normal distributions, Spearman correlation coefficients are reported throughout. Finally, we examined the intercorrelations among the 5 effort-based paradigms and conducted a factor analysis to evaluate whether the paradigms reflect indicators of a unitary underlying construct vs multiple distinguishable constructs.

Results

Descriptive Statistics

Table 1 presents means and *SDs* for the key variables. This outpatient sample was chronically ill and, on average, had moderate levels of symptoms. Mean MCCB scores are typical for this type of sample. Functioning in the area of work was poor, though somewhat better in the areas of independent living and family/social networks.

Negative Symptoms

Because there is currently no consensus on the optimal analytic approach to examine relations between

Table 1. Descriptive Statistics for Schizophrenia Sample (*N* = 94)

Demographic and Clinical Variables	
Sex (% male)	69%
Age	49.1 (11.7)
Education	13.1 (1.8)
Parental Education	13.5 (2.5)
Age at first hospitalization	25.6 (9.45)
Number of hospitalizations	7.6 (8.61)
Antipsychotic medications	
First generation	12%
Second generation	84%
Both types	1%
None	3%
Side effects rating [0–3]	.13 (.31)
Negative symptoms	
CAINS MAP [0–36]	15.6 (7.0)
CAINS Expressive [0–16]	4.9 (4.1)
CAINS total [0–52]	21.0 (9.4)
NSA-16 Global [1–7]	3.5 (1.3)
PANSS Negative [7–49]	16.0 (7.0)
Other symptoms	
PANSS Positive [8–56]	18.5 (7.5)
PANSS Depression [4–28]	7.1 (2.7)
PANSS Disorganized [7–49]	12.6 (4.6)
Motivation, community functioning and neurocognition	
QLS Intrinsic Motivation [0–6]	2.7 (1.2)
RFS work [1–7]	2.6 (1.7)
RFS independent living [1–7]	4.7 (1.6)
RFS family/social [1–7]	4.5 (1.5)
MCCB overall composite	31.6 (12.2)
Self-reported traits	
Behavioral Inhibition Scale [7–28]	20.0 (3.94)
Behavioral Activation Scale [13–52]	39.2 (5.41)
Defeatist Performance Attitudes Scale [15–105]	51.7 (14.25)

Note: CAINS, Clinical Assessment Interview for Negative Symptoms; MAP, Motivation and Pleasure; NSA, Negative Symptom Assessment; PANSS, Positive and Negative Syndrome Scale; QLS, Quality of Life Scale; RFS, Role Functioning Scale; MCCB, MATRICS Consensus Cognitive Battery. The range of possible scores is presented in brackets. MCCB Overall Composite is based on *T*-scores.

negative symptoms and performance on effort tasks,³ we used 3 alternative approaches. First, we used a Repeated Measures (RM)-ANCOVA approach in which reward level was a within-subject factor, CAINS MAP, our primary negative symptom measure, was included as a covariate, and percent hard choices was the dependent variable for each task. Second, we used a categorical approach in which participants were dichotomized based on a median split on CAINS MAP ratings; high (*M* = 21.8; *SD* = 5.0) vs low (*M* = 10.6; *SD* = 3.5). Subgroups were then compared in RM-ANOVAs with reward level as a within-subject factor and percent hard choices was the dependent variable. Third, a correlational approach was used in which each negative symptom measure (CAINS, NSA-16, and PANSS) was correlated with the difference score for each effort task.

ANCOVA Approach

As shown in [table 2](#), the main effect of reward level was significant for all tasks, indicating higher frequencies of hard choices as reward levels increased, and the main effect of the covariate was not significant for all tasks. There were, however, significant interaction effects for the Balloon Effort Task and the EEfRT Probability variable, indicating that higher MAP scores were associated with smaller increases from low to high reward levels on the Balloon Effort Task and with smaller increases from the low to high probability conditions on the EEfRT ([supplementary figure 1](#)).

Categorical Approach

RM-ANOVA comparisons of high vs low negative symptom subgroups are summarized in [table 3](#). The main effect of reward level was again significant for all tasks. There was a significant group effect for the Perceptual Effort Task, indicating that the High negative symptoms subgroup selected hard tasks less frequently than the Low negative symptoms subgroup across reward levels ([figure 1](#)). There was also a trend-level interaction for the Balloon Effort Task, which indicated that the High negative symptoms subgroup tended to select the more difficult response less frequently than the Low negative symptoms subgroup at the highest 2 reward levels, though the percentage of hard choices were similar across groups at lower reward levels ([supplementary figure 2](#)).

Correlational Approach

For correlational analyses, EEfRT Probability difference scores showed significant, though small, correlations with

the CAINS MAP ($r = -.23, P = .04$), CAINS EXP ($r = -.22, P = .04$), CAINS Total ($r = -.27, P = .01$), and PANSS negative symptoms ($r = -.23, P = .04$), as well as a trend-level relationship with NSA-16 ($r = -.21, P = .06$). These correlations indicate that higher negative symptoms were associated with less frequent choices of hard tasks at the high vs the low probability level. For the Balloon Effort Task, there was a significant negative correlation with ratings on CAINS MAP ($r = -.25, P = .04$), indicating that higher negative symptoms were associated with less frequent choices of hard tasks at high vs low reward levels. Aside from a trend-level association between the Perceptual Effort Task and CAINS EXP ($r = -.21, P = .06$), there were no significant or trend level-associations with negative symptoms for other tasks. Full results are presented in [supplementary table 1](#).

Motivation

Results are summarized in [table 4](#). The Quality of Life Scale Intrinsic Motivation index showed significant, small to medium correlations across all effort tasks with only 1 exception, the EEfRT Reward variable.

Community Functioning

Higher scores on the Balloon Effort Task and both EEfRT variables were associated with better work functioning, and higher scores on the EEfRT Probability variable correlated with higher levels of independent living ([table 4](#)).

Self-Reported Traits

For the BIS/BAS scales, there was only 1 significant correlation. Higher BAS scores correlated with higher

Table 2. RM-ANCOVA Results Including Negative Symptoms as a Covariate

	Main Effect: Reward Level	Covariate: Negative Symptoms	Interaction
Deck Choice Effort task	$F(2,150) = 14.73^{***}$	$F(1,75) = .32$	$F(2,150) = 1.42$
Perceptual Effort task	$F(3,237) = 9.18^{***}$	$F(1,79) = 1.05$	$F(3,237) = .34$
Balloon Effort task	$F(4,264) = 15.61^{***}$	$F(1,66) = .44$	$F(4,264) = 3.00^*$
Grip Strength Effort task	$F(2,162) = 28.23^{***}$	$F(1,81) = .01$	$F(2,162) = .72$
EEfRT reward	$F(2,164) = 14.64^{***}$	$F(1,82) = .12$	$F(2,164) = 1.04$
EEfRT probability	$F(1,82) = 16.96^{***}$	$F(1,82) = .06$	$F(1,82) = 5.14^*$

Note: RM, Repeated Measures; EEfRT, Effort Expenditure for Rewards Task.

* $P < .05$; *** $P < .001$

Table 3. RM-ANOVA Results Comparing High vs Low Negative Symptoms (CAINS-MAP) Subgroups

	Main Effect: Reward Level	Main Effect: Group	Interaction
Deck Choice Effort task	$F(2,152) = 48.74^{***}$	$F(1,75) = .04$	$F(2,152) = 1.66$
Perceptual Effort task	$F(3,237) = 41.83^{***}$	$F(1,79) = 4.11^*$	$F(3,237) = .19$
Balloon Effort task	$F(4,264) = 34.41^{***}$	$F(1,66) = .07$	$F(4,264) = 2.28^\dagger$
Grip Strength Effort task	$F(2,164) = 122.93^{***}$	$F(1,82) = .12$	$F(2,164) = .16$
EEfRT reward	$F(2,164) = 56.41^{***}$	$F(1,82) = .12$	$F(2,164) = .55$
EEfRT probability	$F(1,82) = 25.45^{***}$	$F(1,82) = .54$	$F(1,82) = 1.39$

Note: $^\dagger P < .10$; * $P < .05$; *** $P < .001$

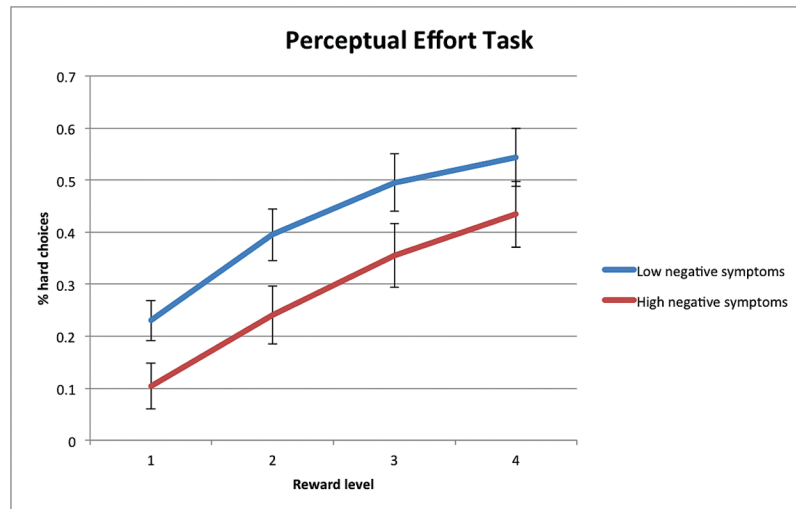


Fig. 1. Percentage of hard choices across reward levels on the Perceptual Effort Task in Low negative symptoms and High negative symptoms subgroups.

Table 4. Correlations Between Effort Tasks and Intrinsic Motivation, Functioning, and Neurocognition

	QLS Intrinsic Motivation ^a	RFS Work ^b	RFS Independent Living ^a	RFS Family/Social ^a	MCCB Composite ^a
Deck Choice Effort task	.26*	.05	.03	.13	.31**
Perceptual Effort task	.25*	.14	.05	.07	.16
Balloon Effort task	.26*	.38**	.04	.11	.39**
Grip Strength Effort task	.23*	.16	.02	-.02	.29**
EEfRT reward	.18	.26*	.04	.13	.25*
EEfRT probability	.34**	.26*	.30**	.19	.29**

Note: ^aZ-test comparisons (2-tailed) of correlation coefficients revealed no significant differences between paradigms.

^bZ-test comparisons (2-tailed) of correlations: Balloon Effort Task > Deck Choice Effort Task.

* $P < .05$; ** $P < .01$.

difference scores on the Perceptual Effort Task ($r = .27$, $P = .02$). There were no significant or trend-level associations with the DPAS.

Neurocognition

All effort tasks showed significant, small to medium correlations with neurocognition on the MCCB, except for the Perceptual Effort Task (table 4). Thus, people with better cognitive performance were more willing to work for higher levels of reward. The neurocognitive subdomains of working memory and speed of processing showed the most consistent relations to the effort tasks (supplementary table 2).

Other Psychiatric Symptoms

None of the effort tasks showed significant correlations with ratings on the PANSS positive and depression/anxiety subscales. PANSS disorganized symptoms showed trend-level correlations only with the Perceptual Effort Task and the EEfRT Probability (r 's = $-.21$, P 's = $.06$).

Clinical and Demographic Variables

There were no significant correlations between effort-based variables and ratings of side effects, and no significant differences on effort-based variables between patients taking first- vs second-generation antipsychotics. There were no correlations involving age at first hospitalization or total hospitalizations, though 1 trend-level correlation was found between total hospitalizations and the EEfRT Reward variable ($r = .25$, $P = .06$). Neither age nor education significantly correlated with any of the effort tasks. There was a significant sex effect only for the Grip Effort Task; men ($M = 11.0$; $SD = 6.4$) had higher difference scores than women ($M = 7.3$; $SD = 7.0$), $t(82) = 2.48$, $P = .02$.

Subjective Valuation of Money

We first compared the patient sample to the matched control sample ($n = 40$; described in Reddy L. R. et al.² this issue) on mean scores for \$10 (patients: $M = 6.1$, $SD = 3.0$; controls: $M = 5.0$, $SD = 2.9$) and \$1000 (patients: $M = 9.0$, $SD = 2.0$; controls: $M = 9.2$, $SD = 1.5$)

using RM-ANOVA. There was a significant main effect of dollar amount, $F(1,132) = 190.03$, $P < .001$, reflecting higher overall valuation of \$1000 vs \$10, a nonsignificant group effect ($P = .30$) and a significant interaction, $F(1,132) = 6.00$, $P = .02$. The interaction indicated that patients valued \$10 more than controls while both groups assigned similar ratings to \$1000.

Among patients, scores on the subjective valuation of money index (ie, valuation of \$1000–\$10) significantly and positively correlated with the EEfRT Reward variable ($r = .27$, $P = .01$). Thus, patients who showed a larger valuation difference between \$1000 vs \$10 were more likely to select difficult tasks for high vs low rewards. There were no other significant relationships with effort tasks, though trend-level correlations in the same direction were found for the Deck Choice Effort Task ($r = .21$, $P = .06$) and the Balloon Effort Task ($r = .21$, $P = .08$). Finally, correlations with negative symptoms were nonsignificant ($r = .13$ for CAINS MAP; $r = -.01$ for CAINS Expression).

Interrelationships Among Effort Tasks

Intercorrelations are shown in [table 5](#). In general, the effort tasks showed significant, medium associations with each other, ranging from about .30 to .45. The main exception to this pattern was the EEfRT Probability variable, which showed generally nonsignificant and lower correlations with the other measures. The pattern was similar when analyses were restricted to participants who validly completed all paradigms ([supplementary table 3](#)).

We examined whether the effort tasks are indicators of a common factor or instead relate to multiple underlying factors. The EEfRT Probability variable was excluded from these analyses due to its conceptual difference from the other variables, which all focused on sensitivity to reward. A maximum likelihood factor analysis revealed a clear 1-factor solution that explained 53.0% of the variance (Scree plot presented in [supplementary figure 3](#)). Factor loadings were uniformly high: Grip Effort Task = .76; Deck Effort Task = .74; EEfRT Reward = .66; Balloon Effort Task = .55; Perceptual Effort Task = .48; and the Goodness-of-fit Test was acceptable, $\chi^2 = 7.48$, $P = .19$. Thus, the 5 effort tasks appear to reflect a

relatively cohesive set of indicators of a unitary underlying construct.

Discussion

This study examined the external validity and correlates of 5 effort-based decision-making paradigms. The paradigms showed generally small to medium relationships to clinical ratings of negative symptoms, functioning, and motivation, though the pattern was more consistent for some measures than others. They also showed medium relations with neurocognitive functioning, but generally small and nonsignificant relations to self-reported traits, other psychiatric symptoms, type of antipsychotic medications, side effects, and subjective valuation of money. The paradigms all showed relatively strong relationships to a single underlying factor. In conjunction with the companion article (Reddy L. R. et al,² this issue), these results provide modest support for the potential usefulness of effort-based decision-making paradigms in clinical trials.

Relations to Clinically Rated Negative Symptoms, Motivation, and Functioning

Novel interventions are under development that are ultimately intended to improve negative symptoms and their associated functional impairments. For the reason, the most critical external variables include clinically rated negative symptoms, intrinsic motivation, and community role functioning.

Regarding negative symptoms, significant associations were found in the expected direction for the EEfRT, Balloon Effort, and Perceptual Effort Tasks. EEfRT Probability showed relations with the CAINS MAP scale (and other negative symptom scales) using 2 of the analytic approaches. These findings converge with 2 of the 3 prior studies using the EEfRT in schizophrenia, which found correlations with either clinical ratings of avolition⁴ or (at a trend level) total negative symptoms.⁷ The Balloon Effort Task showed associations with CAINS MAP assessed both continuously and categorically. A prior study with this paradigm found a high negative symptom subgroup showed diminished effort expenditure compared to healthy controls (though not a low negative symptom subgroup) but no correlations with continuous ratings.⁵ The Perceptual Effort Task significantly distinguished High vs Low CAINS MAP patients, which is interesting since this was the only paradigm that did not distinguish patients from controls in our companion article (Reddy L. R. et al,² this issue). Notably, the relations between these effort tasks and negative symptoms were generally small in magnitude across negative symptom indices.

The most consistent external correlate across the effort tasks was clinical ratings of intrinsic motivation. Small to medium correlations were found in the expected direction

Table 5. Intercorrelations Among the 5 Effort Tasks

	1	2	3	4	5
1. Deck Choice Effort task	—				
2. Perceptual Effort task	.38**	—			
3. Balloon Effort task	.38**	.14	—		
4. Grip Strength Effort task	.43**	.35**	.47**	—	
5. EEfRT reward	.42**	.46**	.48**	.34**	—
6. EEfRT probability	.20	.26*	.30*	.13	.20

Note: * $P < .05$; ** $P < .01$.

for each of the effort tasks (though for the EEfRT task it was only the Probability, not the Reward, index). These associations with clinically relevant motivational disturbances provide some support for the external validity of all of the effort paradigms. Clinically rated role functioning was related to the EEfRT and Balloon Tasks. EEfRT Reward and Probability both correlated with work functioning, while Probability also correlated with independent living. A prior study that found performance on a different version of the EEfRT correlated with work and community functioning,⁴ though another found no associations with subjective quality of life.¹⁰ Performance on Balloon Effort also correlated with work functioning. There is, thus, some consistency across studies in that different physical effort tasks show associations with work functioning.

Overall, the Balloon Effort Task and EEfRT showed the broadest patterns of relation to the primary external validity measures, though the relations were small to medium. In the context of the generally small relationships found in this study, it is worth noting that there may be some trade-offs when using neuroscience-based effort tasks. There are clear advantages to selecting measures whose neural substrates have been extensively studied for treatment development. However, there are complex intervening steps on the causal pathway between these relatively discrete neural processes and clinically rated symptoms and outcomes, which would serve to diminish direct correlations with external variables.³⁵ One might also question the use of clinical ratings of negative symptoms as the “gold standard” in these analyses, since these ratings do not directly address effort and are subject to confounding factors such as memory biases, insight, and willingness to disclose.³ It remains to be determined why the relations between effort task performance and experiential negative symptoms were smaller than anticipated.

Correlations With Other Variables

We examined a variety of other variables to determine the degree to which effort-based tasks correlate with key features of schizophrenia. Our interest in neurocognition stems from findings that negative symptoms are often associated with neurocognitive impairment,³⁶ and the proposal that neurocognitive processes may play a role in impaired effort-based decision making.^{5,29} Performance on each of the paradigms, except for Perceptual Effort, showed medium correlations with neurocognition, indicating that they were meaningfully related to neurocognition but clearly not redundant with it. Poorer neurocognition was associated with smaller increases in effort allocation from lower to higher reward (and probability) levels. A prior study reported a similar correlation with the Balloon Effort Task,⁵ although 2 studies using grip effort tasks did not.^{9,10} The current results are consistent with proposals that cognitive impairments contribute

to aberrant decision making by impacting effort computations and/or difficulty maintaining representations of reward value.^{9,10}

Regarding other correlates, performance on the effort paradigms was generally not associated with self-reported traits related to BIS/BAS sensitivity or defeatist performance attitudes, which is consistent with 2 prior studies.^{4,10} Second, as in most prior studies,^{5,6,10} the effort tasks were unrelated to other psychiatric symptoms. Third, no associations were seen with type of antipsychotic medications or with extrapyramidal side effects. Although D2 antagonists can reduce willingness to work for rewards in normal rats,^{37–39} all prior studies have failed to find a link between antipsychotic medications and diminished willingness to exert effort on decision-making tasks in schizophrenia. Fourth, performance was generally unrelated to other clinical and demographic characteristics. The 1 exception was a sex effect for the Grip Effort task indicating that men more frequently chose hard grip tasks than women.

This study also considered the impact of subjective monetary reward valuation. One might be concerned that patients value money differently than healthy individuals due to the relatively lower socio-economic status often associated with clinical samples. Two prior studies reported no patient vs control differences in value ratings for a single⁶ or multiple¹⁰ amounts of money. We found that, although patients and controls reported similar valuation for a relatively large amount (\$1000), patients reported higher valuation of a relatively small amount (\$10). Taken together, these findings of intact or even higher valuation of money argue against the possibility that diminished effort exertion in schizophrenia is due to decreased valuation of monetary rewards. The patients' pattern is consistent with a preclinical model relevant to schizophrenia in which mice that over-express striatal dopamine D2 receptors demonstrate intact ‘hedonic’ valuation of reward yet reduced willingness to expend effort for reward.⁴⁰

Finally, we explored whether the 5 effort paradigms reflect a single or multiple underlying constructs. Results suggested a unitary construct with each paradigm loading strongly on this dimension. These conceptually related paradigms appear to cohesively hang together, which supports their convergent validity in schizophrenia. Pending replication, this result may have practical implications for assessment (eg, a single summary score may be appropriate for a battery of effort tasks) and for statistical modeling (eg, a single latent trait may load on multiple effort-based decision-making tasks).

Additional factors should be considered regarding the current findings. First, almost all patients were taking antipsychotic medications. That said, the clinical reality remains that most individuals with schizophrenia are on such medications, and all drugs currently under development for negative symptoms are intended

for use as adjunctive medications. Second, this sample included a larger proportion of participants from Veterans Administration clinics than prior studies of effort-based tasks and may thus have a relatively smaller proportion with poor premorbid functioning, who often have the highest levels of persistent negative symptoms. Third, this was a chronically ill sample and additional studies in recent-onset and unmedicated at-risk samples are needed to clarify the course and potential etiological significance of effort-based decision-making impairments. Fourth, our measures of community functioning were based solely on patient reports without information from informants, which could limit their validity.³³

Summary and Conclusion

The current study, along with a handful of prior studies, provides converging evidence for the validity and potential usefulness of effort-based decision-making tasks for clinical trials in schizophrenia. The performance of the 5 paradigms examined in the current and companion articles (Reddy L. R. et al,² this issue) on the criteria for clinical trial outcome measures is summarized in [table 6](#).

At this juncture, the EEfRT Probability index and Balloon Effort Task show the broadest evidence for external validity, as well as good discrimination between patients vs controls, promising test-retest reliability, and good tolerability. However, these paradigms had notable limitations. These tasks had the longest administration durations, which raises practical concerns regarding time constraints in clinical trials. They are also less “process pure” than the others. For both paradigms, hard vs easy tasks differed in effort requirements, as well as the associated time intervals between

choice selection and reward receipt, which involves concomitant temporal discounting. Furthermore, the EEfRT task also incorporated a probability manipulation, which makes additional computational demands, and the Balloon Effort Task had a relatively high proportion of inflexible responders, which present a practical obstacle for its ability to show treatment effects in clinical trials. Further paradigm development efforts could examine whether shorter versions show adequate sensitivity, modifications to task parameters can reduce the number of inflexible responders, and equating trial length for hard and easy tasks (to address conflation with temporal discounting) impacts the performance profiles. Overall, these 2 paradigms appear to be comparatively further along the pathway to suitability for use in clinical trials.

The Grip Effort Task, Deck Selection Task, and Perceptual Effort Tasks showed more mixed profiles. These tasks showed less consistent external validity, though each task significantly correlated with the intrinsic motivation index. Furthermore, the Perceptual Effort Task was sensitive to high vs low negative symptom levels but did not discriminate between patients vs controls. Regarding psychometric properties, these paradigms generally performed below typical benchmarks for test-retest reliability and suitability for use as repeated measures (Reddy L. R. et al,² this issue). However, all 3 showed sufficiently promising psychometric properties to warrant further consideration, and they have the benefit of being relatively brief and well tolerated. In addition, each involves similar time intervals for hard vs easy trials and the Grip and Perceptual Effort tasks incorporate individual titration procedures, which minimize confounds associated with temporal discounting and psychomotor slowing, respectively. Thus, these 3 newly developed paradigms may be useful tools with further development.

Table 6. Criteria for Measures for Use in Clinical Trials

	Deck Choice Effort Task	Perceptual Effort Task	Balloon Effort Task	Grip Effort Task	EEfRT Probability	EEfRT Reward
Psychometrics						
Patient vs control difference	✓	O	✓	✓	✓	X
Test-retest reliability	X	✓	X	X	X	✓
Utility as a repeated measure	X	X	✓	X	✓	✓
Feasibility						
Tolerability	X	✓	✓	✓	✓	✓
Duration	✓	✓	O	✓	O	O
Inflexible responders	X	✓	O	✓	✓	✓
Validity						
Negative symptoms	O	✓	✓	O	✓	O
Intrinsic motivation	X	X	X	X	✓	O
Community functioning	O	O	✓	O	✓	X

Note: ✓ = acceptable; X = promising; O = poor; benchmarks for psychometrics: test-retest reliability = ICC > .70; utility as a repeated measure = $d < .20$, tolerability > 5.0, duration < 20 min.

Supplementary Material

Supplementary material is available at <http://schizophreniabulletin.oxfordjournals.org>.

Funding

This project was an investigator-initiated research collaboration funded by Amgen, Inc.

Acknowledgments

Dr Barch has been a paid consultant for Roche, Amgen, Pfizer and Takeda, and has received research funds from Pfizer and DSP. Dr Buchanan: DSMB member: Pfizer; Consultant: AbbVie, Amgen, EnVivo, Roche, and Takeda. Dr Dunayevich is a full-time employee of Amgen. Dr Gold receives royalty payments from the BACS, and has been a paid consultant for Lundbeck, Roche, Amgen, Pfizer, Merck, AstraZeneca, Solvay, and GlaxoSmithKline. Dr Green reports having been a consultant to AbbVie, DSP, Forum, and Takeda, he is a member of the scientific board for Mnemosyne, and he received research funds from Amgen and Forum. Dr Marder reports having been a consultant to Abbott, Roche, Genentech, Otsuka, Bristol Meyers Squibb, Pfizer, Lundbeck, and Boehringer Ingelheim, and he has received research funds from Amgen, Psychogenics, and Sunovion.

References

- Young JW, Markou A. Translational rodent paradigms to investigate neuromechanisms underlying behaviors relevant to amotivation and altered reward processing in schizophrenia. *Schizophr Bull*.
- Reddy LR, Horan WP, Barch D, et al. Effort-based decision making paradigms for clinical trials in schizophrenia: part 1 - psychometric characteristics of five paradigms. *Schizophr Bull*.
- Green MF, Horan WP, Barch DM, Gold JM. Effort-based decision making: a novel approach for assessing motivation in schizophrenia. *Schizophr Bull*.
- Barch DM, Treadway MT, Schoen N. Effort, anhedonia, and function in schizophrenia: reduced effort allocation predicts amotivation and functional impairment. *J Abnorm Psychol*. 2014;123:387–397.
- Gold JM, Strauss GP, Waltz JA, Robinson BM, Brown JK, Frank MJ. Negative symptoms of schizophrenia are associated with abnormal effort-cost computations. *Biol Psychiatry*. 2013;74:130–136.
- Hartmann MN, Hager OM, Reimann AV, et al. Apathy but not diminished expression in schizophrenia is associated with discounting of monetary rewards by physical effort. *Schizophr Bull*. 2015;41:503–512.
- Treadway MT, Peterman JS, Zald DH, Park S. Impaired effort allocation in patients with schizophrenia. *Schizophr Res*. 2015;161:382–385.
- Wolf DH, Satterthwaite TD, Kantrowitz JJ, et al. Amotivation in schizophrenia: integrated assessment with behavioral, clinical, and imaging measures. *Schizophr Bull*. 2014;40:1328–1337.
- Docx L, de la Asuncion J, Sabbe B, et al. Effort discounting and its association with negative symptoms in schizophrenia. *Cogn Neuropsychiatry*. 2015;20:172–185.
- Fervaha G, Graff-Guerrero A, Zakzanis KK, Foussias G, Agid O, Remington G. Incentive motivation deficits in schizophrenia reflect effort computation impairments during cost-benefit decision-making. *J Psychiatr Res*. 2013;47:1590–1596.
- Gold JM, Kool W, Botvinick MM, Hubzin L, August S, Waltz JA. Cognitive effort avoidance and detection in people with schizophrenia. *Cogn, Affect Behav Neurosci*. 2015;15:145–154.
- First MB, Gibbon M, Spitzer RL, Williams JBW, Benjamin L. *Structured Clinical Interview for DSM-IV Axis II Personality Disorders*. New York, NY: Biometrics Research Department, New York State Psychiatric Institute; 1996.
- First MB, Spitzer RL, Gibbon M, Williams JBW. *Structured Clinical Interview for DSM-IV Axis I Disorders - Patient Edition*. New York, NY: Biometrics Research Department, New York State Psychiatric Institute; 1997.
- Treadway MT, Buckholz JW, Schwartzman AN, Lambert WE, Zald DH. Worth the ‘EEfRT’? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PLoS One*. 2009;4:e6598.
- Kring AM, Gur RE, Blanchard JJ, Horan WP, Reise SP. The Clinical Assessment Interview for Negative Symptoms (CAINS): final development and validation. *Am J Psychiatry*. 2013;170:165–172.
- Axelrod BN, Goldman RS, Alphas LD. Validation of the 16-item Negative Symptom Assessment. *J Psychiatr Res*. 1993;27:253–258.
- Kay SR, Fiszbein, A., Opler, L.A. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull*. 1987;13:261–271.
- Heinrichs DW, Hanlon TE, Carpenter WT. The quality of life scale: an instrument for rating the schizophrenic deficit syndrome. *Schizophr Bull*. 1984;10:388–398.
- Nakagami E, Xie B, Hoe M, Brekke JS. Intrinsic motivation, neurocognition and psychosocial functioning in schizophrenia: testing mediator and moderator effects. *Schizophr Res*. 2008;105:95–104.
- Choi J, Choi KH, Felice Reddy L, Fiszdon JM. Measuring motivation in schizophrenia: is a general state of motivation necessary for task-specific motivation? *Schizophr Res*. 2014;153:209–213.
- Bell MD, Choi KH, Dyer C, Wexler BE. Benefits of cognitive remediation and supported employment for schizophrenia patients with poor community functioning. *Psychiatr Serv*. 2014;65:469–475.
- Gard DE, Fisher M, Garrett C, Genevsky A, Vinogradov S. Motivation and its relationship to neurocognition, social cognition, and functional outcome in schizophrenia. *Schizophr Res*. 2009;115:74–81.
- McPheeters HL. Statewide mental health outcome evaluation: a perspective of two southern states. *Community Ment Health J*. 1984;20:44–55.
- Carver CS, White TL. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *J Pers Soc Psychol*. 1994;67:319–333.
- Grant PM, Beck AT. Defeatist beliefs as a mediator of cognitive impairment, negative symptoms, and functioning in schizophrenia. *Schizophr Bull*. 2009;35:798–806.

26. Nuechterlein KH, Green MF, Kern RS, et al. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *Am J Psychiatry*. 2008;165:203–213.
27. Harvey PD, Koren D, Reichenberg A, Bowie CR. Negative symptoms and cognitive deficits: what is the nature of their relationship? *Schizophr Bull*. 2006;32:250–258.
28. Strauss GP, Morra LF, Sullivan SK, Gold JM. The role of low cognitive effort and negative symptoms in neuropsychological impairment in schizophrenia. *Neuropsychology*. 2015;29:282–291.
29. Strauss GP, Waltz JA, Gold JM. A review of reward processing and motivational impairment in schizophrenia. *Schizophr Bull*. 2014;40(Suppl 2):S107–S116.
30. Meyer EC, Carrión RE, Cornblatt BA, et al; NAPLS group. The relationship of neurocognition and negative symptoms to social and role functioning over time in individuals at clinical high risk in the first phase of the North American Prodrome Longitudinal Study. *Schizophr Bull*. 2014;40:1452–1461.
31. Bratti IM, Kane JM, Marder SR. Chronic restlessness with antipsychotics. *Am J Psychiatry*. 2007;164:1648–1654.
32. Goldstein RZ, Tomasi D, Alia-Klein N, et al. Subjective sensitivity to monetary gradients is associated with frontolimbic activation to reward in cocaine abusers. *Drug Alcohol Depend*. 2007;87:233–240.
33. Martin-Soelch C, Chevalley AF, Kunig AF, et al. Changes in reward-induced brain activation in opiate addicts. *Eur J Neurosci*. 2001;14:1360–1368.
34. Bender R, Lange S. Adjusting for multiple testing - when and how? *J Clin Epidemiol*. 2001;54:343–349.
35. Green MF, Helleman G, Horan WP, Lee J, Wynn JK. From perception to functional outcome in schizophrenia: modeling the role of ability and motivation. *Arch Gen Psychiatry*. 2012;69:1216–1224.
36. Ventura J, Wood RC, Helleman GS. Symptom domains and neurocognitive functioning can help differentiate social cognitive processes in schizophrenia: a meta-analysis. *Schizophr Bull*. 2013;39:102–111.
37. Randall PA, Lee CA, Podurriel SJ, et al. Bupropion increases selection of high effort activity in rats tested on a progressive ratio/chow feeding choice procedure: implications for treatment of effort-related motivational symptoms. *Int J Neuropsychopharmacol*. 2014;18.
38. Randall PA, Pardo M, Nunes EJ, et al. Dopaminergic modulation of effort-related choice behavior as assessed by a progressive ratio chow feeding choice task: pharmacological studies and the role of individual differences. *PLoS One*. 2012;7:e47934.
39. Salamone JD, Correa M, Nunes EJ, Randall PA, Pardo M. The behavioral pharmacology of effort-related choice behavior: dopamine, adenosine and beyond. *J Exp Anal Behav*. 2012;97:125–146.
40. Ward RD, Simpson EH, Richards VL, et al. Dissociation of hedonic reaction to reward and incentive motivation in an animal model of the negative symptoms of schizophrenia. *Neuropsychopharmacology*. 2012;37:1699–1707.