Philip Jackson and Philip Pettit have argued that plausible assumptions in the philosophy of language entail that expressivism collapses into subjectivism. Though there have been responses to their argument, I think the responses have not adequately diagnosed the real weakness in it. I suggest my own diagnosis, and defend expressivism as a viable theory distinct from subjectivism.

I. INTRODUCTION

In meta-ethics, it is common to distinguish two related theories, one generally called ‘subjectivism’, the other either ‘emotivism’ or ‘expressivism’. Subjectivism is the view that an ethical sentence of the form

\[ x \text{ is good} \]

means the same as a sentence reporting the speaker’s attitude, in this case, one of the form

\[ \text{I approve of } x. \]

Expressivism, on the other hand, is the view that an ethical sentence expresses, but does not report, the speaker’s attitude, and therefore that (1) is equivalent not to (2), but rather to something like the exclamation

\[ \text{Hooray for } x! \]

or perhaps the prescription

\[ \text{Please, everyone, do what you can to promote } x. \]

Expressivists generally claim that (1), like (3) and (4), has no truth-conditions. That is, they claim that unlike reports of one’s attitudes, ethical sentences express one’s attitudes in such a way that those sentences are neither true nor false. The difference between these two views is thought to be important; subjectivism is supposed to be
open to serious objections which do not affect expressivism. But Frank Jackson and Philip Pettit (henceforth J&P) have argued that certain plausible assumptions in the philosophy of language entail that expressivism is untenable – more precisely, that expressivism collapses into subjectivism.

Relying on a claim in the philosophy of language which they trace back to Locke, J&P argue that since expressivists maintain that an ethical sentence expresses the speaker’s attitude, they must admit that an ethical sentence has truth-conditions – i.e., it is true if and only if the speaker really has the attitude in question, and is otherwise false. So (1) is, after all, equivalent to (2), at least with respect to truth-conditions. The argument is as follows.

According to the Lockean philosophy of language, a sentence \( s \) gets its meaning from a convention, or agreement, to use \( s \) when we believe that a certain set of conditions \( c \) obtains and we think that circumstances are right for expressing, or communicating the content of, that belief. Thus, for example, a sentence of the form

\[
\text{5. } x \text{ is square}
\]

has the meaning it has because

\[
\text{5a. We have agreed to use (5) when we believe that } x \text{ is square and that circumstances are right for communicating this fact (i.e., for expressing, or communicating the content of, our belief that } x \text{ is square).}
\]

Applying this philosophy of language to the expressivist view that we use (i) to express our approval of \( x \), we get:

a. We have agreed to use (i) when we believe that we approve of \( x \) and think that the circumstances are right for expressing that belief.

I take (a) to be equivalent to what J&P express as ‘(B2) belief claim (good): we have agreed to use “\( x \) is good” when we believe that we approve of \( x \) and that conditions are right for communicating this fact’ (LEC, p. 89). My main reason for changing (B2) is to make clearer that by ‘this fact’ J&P mean the fact that \( x \) is square, not the fact that we believe that \( x \) is square. But I have another reason as well: my formulation will make it simpler to highlight part of the difference between subjectivism and expressivism which will emerge in §II below. I hope I am correct in thinking that I have not prejudiced any substantive issues by using my formulation.

J&P claim that

b. If we have agreed to use (i) when we believe that we approve of \( x \) and think that circumstances are right for expressing that belief, then we have agreed to use (i) to report that we approve of \( x \).

This is a paraphrase and particular application of the essentials of J&P’s ‘(A2) Locke’s claim (good): what it would be to use “\( x \) is good” to stand for \( x \)’s being such


© The Editors of The Philosophical Quarterly, 2004
and such is to agree to use “x is good” when we believe that x is such and such, and that conditions are right for communicating this fact’ (LEC, p. 89). I use (b), rather than J&P’s own formulation, because I think it makes the structure of their argument more clear.

From (a) and (b) it follows that

c. We have agreed to use (1) to report that we approve of x.

Since we have also agreed to use (2) to report that we approve of x, it follows that according to expressivism, (1) is equivalent to (2); and so expressivism collapses into subjectivism.

Michael Smith and Daniel Stoljar (henceforth S&S) defend expressivism against J&P’s argument as presented in PE. Their main point is that J&P overlook a distinction which undermines their argument, the distinction between agreeing to use s for c’s obtaining and agreeing merely to use s when c obtains – or more briefly, between a for-agreement and a when-agreement. For example, (5a) describes a for-agreement; we have agreed to use (5) not only when we believe that x is square, but also for x’s being square, that is, with the intention of reporting that x is square. S&S claim that in some cases (i.e., for some s) we agree to use s when we believe that c obtains, but not for c’s obtaining, that is, not to report that c obtains. In particular, they claim that according to expressivism, we have agreed to use (1) when we believe that we approve of x, but not to report our approval of x; so (1) is not equivalent to (2).

I think that the best way to apply S&S’s distinction to my version (based on J&P’s later article) of J&P’s argument is to invoke a related, or parallel, distinction that S&S (p. 83) draw between weak and strong expression of a belief. A use of s weakly expresses the speaker’s belief that c obtains if and only if we have agreed to use s when we believe that c obtains. A use of s strongly expresses the speaker’s belief that c obtains if and only if we have agreed to use s when (i) we believe that c obtains, and (ii) we intend, by using s, to report the content of that belief (i.e., to report that c obtains). S&S claim, in effect, that (a) entails only that we use (1) to express weakly our belief that we approve of x, whereas (b) is true only if it is about strongly expressing that belief; therefore the argument is unsound, and appears sound only because it equivocates with respect to the two senses of ‘express a belief’.

J&P have responded, summing up their disagreement with S&S as follows: ‘The issue is whether agreeing to use words when you believe such and such and that conditions are right for communicating this fact is to agree to use your words for such and such. We follow Locke in saying that it is’ (LEC, p. 88). Thus J&P reject S&S’s distinction between weak and strong expression of belief: to express a belief is to report the content of the belief. So (a) entails, in effect, that we use (1) to express strongly our belief that we approve of x; and so there is no equivocation in the argument. As for the distinction between when-agreements and for-agreements, J&P might admit that in some contexts it is important; but in the context of agreeing to use a sentence to express a belief, the distinction is irrelevant. Thus, given that the expressivist is committed to (a), S&S’s distinction does not save expressivism from

1 M. Smith and D. Stoljar, ‘Is There a Lockean Argument against Expressivism?’, Analysis, 63 (2003), pp. 76–86.
collapsing into subjectivism. I am inclined to agree with J&P on this point; for if s is conventionally used to express the belief that c obtains, then it is unclear what one’s intention would be in using s, if it were not to report that c obtains.

However, I shall try to show that expressivists are not committed to the view that we use (i) to express our belief that we approve of x. I shall argue that according to expressivism, we use (i) to express our approval of x. Unlike S&S, I shall argue that according to expressivism, the agreement governing (i) is not to use (i) when we believe that we approve of x, but simply to use (i) when we approve of x. This kind of agreement is not a for-agreement, and therefore something like S&S’s distinction is relevant. But my criticism of J&P’s argument is not essentially dependent on S&S’s criticism; it is simply that even according to the Lockean philosophy of language, expressivists are not committed to (a).4

II. BELIEF AND APPROVAL

I shall argue that (a) does not correctly represent expressivism, and should be replaced by

a’. We have agreed to use (i) when we approve of x and think that circumstances are right for expressing that approval.

That is, we should replace ‘we believe that we approve of x’ in (a) by ‘we approve of x’ (and ‘belief’ by ‘approval’ accordingly). My point is that (a) represents a false premise in J&P’s argument; we could solve this problem by substituting (a’) for (a), but then the resulting argument would be invalid.

In this section, I argue that from J&P’s Lockean perspective, expressivists, in effect, do accept (a’) rather than (a); and in the next section, I shall argue that expressivists can consistently accept (a’) and not (a).

I can show that (a’), rather than (a), correctly represents expressivism, by exploring the following recognizably expressivist account, within a Lockean framework, of how (i) gets its meaning. (I am not endorsing this account, just as throughout this paper I am not endorsing expressivism. I am merely presenting the expressivist point of view in such a way as to show that it is not vulnerable to J&P’s argument.)

Many of us do not realize that ethical judgements are expressions of our attitudes – for example, that we use (i) to express our approval of x. This is because we tend to project our approval of x onto x itself, and imagine (with varying degrees of clarity) that x has an intrinsic normative property, goodness, a property which we detect by means of a cognitive faculty of moral intuition. We therefore think that we have agreed to use (i) when we believe, on the basis of moral intuition, that x has the intrinsic property of goodness. But, the expressivist holds, in fact we have no such cognitive faculty, and there is no such intrinsic normative property. Still, we have made a useable agreement; for the terms in which we tend to express it, though

4 Smith and Stoljar (pp. 79 fn. 5, 81 fn. 8) mention something like this as a possible response to Jackson and Pettit, but do not adopt it or develop it.

© The Editors of The Philosophical Quarterly, 2004
inaccurate as descriptions, nevertheless do refer to actual phenomena. What we are referring to as ‘believing, on the basis of intuition, that $x$ has the property of goodness’ is really just approving of $x$; and what we are referring to as ‘the intrinsic property of goodness’ is really just the relational (or indexical) property of being approved by us, or of evoking, or tending to evoke, our approval. Thus we have agreed to use (i) when we approve of $x$, to express our approval, though we may not have realized that this was what we were doing.

What emerges is that according to expressivism, to say that, for example, Mary believes that $x$ is good is to say that Mary approves of $x$, not (as the subjectivist would have it) that Mary believes that she approves of $x$. Thus the average person (or an intuitionist) might say

1a. We have agreed to use (i) when we believe that $x$ is good and think that the circumstances are right for expressing that belief.

A subjectivist would analyse, or rationally reconstruct, (1a) as

a. We have agreed to use (i) when we believe that we approve of $x$ and think that the circumstances are right for expressing that belief.

But an expressivist would rationally reconstruct (1a) as

a’. We have agreed to use (i) when we approve of $x$ and think that the conditions are right for expressing that approval.

It will perhaps be more evident that (a), in J&P’s argument, should be replaced by (a’) if I compare various things that expressivists would say about

1. $x$ is good

with what they (or anyone) would say about

5. $x$ is square.

We can establish a kind of proportion: believing that $x$ is square is to (5) as approving of $x$ (as opposed to believing that one approves of $x$) is to (1). For example:

A. If Mary says that $x$ is square, then by saying so, she implies (pragmatically) that she believes that $x$ is square, though (5) itself does not (semantically) entail that the speaker or anyone else believes that $x$ is square. Analogously, according to expressivism, if Mary says that $x$ is good, then by saying so, she implies (pragmatically) that she approves of $x$; but (1) itself does not (semantically) entail that the speaker or anyone else approves of $x$.5

B. If John challenges Mary, saying ‘Do you really believe that $x$ is square?’, then Mary, if she takes the challenge seriously, will reconsider the reasons for and against believing that $x$ is square, and then answer ‘Yes’ or ‘No’, depending on whether this reconsideration has led her to believe that $x$ is square. If John challenges Mary, saying ‘Do you really believe that $x$ is good?’, then Mary, if

she takes the challenge seriously, will reconsider the reasons for and against approving of \( x \), and then answer ‘Yes’ or ‘No’, depending on whether this reconsideration has led her to approve of \( x \).

C. If John, in both cases, were a psychologist in a therapy session trying to help Mary get in touch with her thoughts and feelings, then Mary would probably respond instead by introspecting and examining her own behaviour, in the one case to determine whether she really believes that \( x \) is square, and in the other to determine whether she really approves of \( x \).

D. If Mary says that \( x \) is square and John responds ‘That is true’, John is indicating that he shares Mary’s belief that \( x \) is square. Analogously, if Mary says that \( x \) is good, and John responds ‘That is true’, John is indicating that he shares Mary’s approval of \( x \).

These examples illustrate the fact that according to expressivism, Mary’s belief that \( x \) is square corresponds not to her belief that she approves of \( x \), but to her approval of \( x \). So it stands to reason that according to expressivism, believing that \( x \) is square in the agreement governing (5) corresponds to approving of \( x \) in the agreement governing (1). J&P tell us that the agreement governing (5) is, in effect,

5a. We have agreed to use \( ^{(5)} \) when we believe that \( x \) is square and think that circumstances are right for expressing that belief.

Making the appropriate substitutions, we get the result that the agreement governing (1), according to expressivism, must be \( (a') \), not \( (a) \). Therefore J&P’s argument is unsound.

III. ANSWERS TO POSSIBLE OBJECTIONS

In response to the above criticism, J&P might press the Lockean point that we do not express our approval of \( x \) if we do not believe that we approve of \( x \); and they could then claim that truth-conditions are introduced by the fact that believing that we approve of \( x \) is generally involved when we use (1). The basis of this objection would be Locke’s argument that it makes sense to agree to use \( s \) when \( c \) obtains only if we can sometimes know that \( c \) obtains, and that adhering to an agreement to use \( s \) when \( c \) obtains means using \( s \) when we believe that \( c \) obtains.

My answer is, first of all, that the point that we do not express our approval of \( x \) unless we believe that we approve of \( x \), though correct, is analogous to the point that we do not express our belief that \( x \) is square unless we believe that we believe that \( x \) is square. But the agreement governing (5) does not mention our believing that we believe that \( x \) is square. So this point does not require that we mention believing that we approve of \( x \) in the agreement governing (1); that is, it does not imply that (a), as opposed to \( (a') \), belongs in J&P’s argument.

Secondly, even though we use (1) only when we believe that we approve of \( x \), it does not follow that our having that belief confers truth-conditions on (1); for it seems reasonable to say that it is only having a belief which we intend to express by a use of \( s \) that confers truth-conditions on \( s \) (see PE, pp. 246–7).
Thus, for example, the belief which we intend to express by a use of \( (\varphi) \) is the belief that \( x \) is square; and that is why a use of \( (\varphi) \) is true if and only if that belief is true – i.e., if and only if \( x \) is square. But consider the belief \( b(r) \) – that the circumstances are right for expressing the belief that \( x \) is square. \( b(r) \) is generally present when someone utters \( (\varphi) \). But \( b(r) \) is irrelevant to the truth-conditions of \( (\varphi) \); that is, it does not matter to the truth-value of a given use of \( (\varphi) \) whether the circumstances were right for expressing the belief that \( x \) is square. This is because we do not intend to express \( b(r) \) by using \( (\varphi) \). Or consider the belief, or meta-belief, \( b(b) \), that we believe that \( x \) is square. Generally when we use \( (\varphi) \) to express our belief that \( x \) is square, we believe that we believe that \( x \) is square, that is, we also have \( b(b) \). But what we intend to express by a use of \( (\varphi) \) is our belief that \( x \) is square, not \( b(b) \). (We would express \( b(b) \) by reporting the first-order belief, saying ‘I believe that \( x \) is square’.) That is why it is the truth of the belief that \( x \) is square that is necessary and sufficient for the truth of a use of \( (\varphi) \); whereas the truth or falsity of \( b(b) \) is irrelevant to the truth or falsity of a use of \( (\varphi) \). Similarly, according to expressivism, what we intend to express by a use of \( (\iota) \) is our approval of \( x \), not our belief that we approve of \( x \). (We would express the latter belief by reporting our approval, saying ‘I approve of \( x \’ \).) And that is why the truth or falsity of the belief that we approve of \( x \) is irrelevant to the truth-value of \( (\iota) \), and the fact that we believe that we approve of \( x \) is irrelevant to whether \( (\iota) \) has truth-conditions.

As for the Lockean basis, I would say that Locke’s requirement of cognition of the relevant conditions \( c \) is satisfied in \( (a') \) because our approving of \( x \) constitutes our awareness, or cognition, of \( x \’ \)s goodness, that is, of \( x \’ \)s evoking our approval. Belief that we approve of \( x \) is cognition not of \( x \’ \)s goodness, but of our thinking \( x \) good; and therefore, as I have said above, it is irrelevant to the question of whether \( (\iota) \) has truth-conditions.

J&P present an argument that might be thought (mistakenly) to constitute an objection to some of what I have said. Expressivists often appeal to the distinction between expressing and reporting a belief, in an attempt to clarify expressivism and distinguish it from subjectivism. (I made use of the distinction above, in answering the objection raised in the present section.) J&P, however, argue that ‘although there is an important difference between reporting and expressing a belief, it is plausibly a difference in what is reported. It is not a difference between reporting something and not reporting at all’ (PE, p. 245). That is, an expression of belief is itself a report of what is believed, and therefore true or false. Thus, for example, the expression of the belief that snow is white, though not a report of that belief, nevertheless is a report – i.e., of snow’s being white – and so is itself true or false. They conclude that the distinction between reporting and expressing a belief does not serve to support or clarify the expressivist position that a use of \( (\iota) \) expresses, but does not report, the speaker’s approval of \( x \), and therefore has no truth-conditions.

J&P are right. Distinguishing between reporting and merely expressing a belief does not prove that expressions of approval have no truth-conditions; nor does a sentence expressing a belief provide a model of a sentence without truth-conditions, after which to pattern expressions of approval. On the contrary, the case of belief shows that a mere expression of a state of mind can have truth-conditions. But
though we have here neither proof nor model (and we have an example which counters the generalization that mere expressions of a state of mind have no truth-conditions), nevertheless the analogy between reporting vs expressing belief and reporting vs expressing approval still holds; and I think that this analogy can be used to clarify expressivism. The disanalogy which J&P point out does not undermine the distinction between reporting and merely expressing a state of mind; nor does it detract from the heuristic value of comparing expression of belief with expression of approval. The disanalogy is due simply to the difference between belief and approval themselves. The nature of belief, which one might characterize as epistemic endorsement of a proposition, is such that the most natural linguistic way to express (but not report) a belief is to assert the proposition believed; and an assertion of a proposition is true or false. The nature of approval, which one might characterize as affective endorsement of a person, action, or state of affairs (either tokens or types), is such that the way to express linguistically (but not report) approval of $x$ is not to assert a proposition, but rather to exclaim ‘Hooray for $x$’ or else simply say ‘$x$ is good’; neither of these two modes of expression is true or false.

IV. CONCLUSION

I have argued that (a) in J&P’s argument must be replaced by (a’). This replacement would render their argument invalid. They might regain validity by replacing (b) with

(b’). If we have agreed to use (i) when we approve of $x$ and think that circumstances are right for expressing that approval, then we have agreed to use (i) to report that we approve of $x$.

However, unlike (b), (b’) is not plausible. It is plausible to say that to use a sentence $s$ with the intention of expressing my belief that $c$ obtains is to report that $c$ obtains; but I may very well use $s$ in this way without thereby reporting that I believe that $c$ obtains. Similarly, it seems that I may very well use a sentence with the intention of expressing my approval of $x$ without thereby reporting that I approve of $x$. Therefore, using S&S’s terms, we may say that according to expressivism, the agreement governing (i), given in (a’), is only a when-agreement; whereas according to subjectivism, the agreement governing (i), given in (a), is a for-agreement.

I conclude that J&P have not successfully shown that expressivism collapses into subjectivism.6

Bar-Ilan University

---

6 I would like to thank David Widerker, Jerome I. Gellman and Charlotte Katzoff, as well as two anonymous referees, for help with this paper.

© The Editors of The Philosophical Quarterly, 2004
Copyright of Philosophical Quarterly is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.