The Society for Descriptive Psychology
www.sdp.org

Galaxies at 8:30

Peter G. Ossorio
1988

Transcription of a talk presented to The Society for Descriptive Psychology Tenth Annual Conference, October 9, 1988. Originally published as LRI Report No. 42a.

The last three years, I've talked about appraisal and I've talked about significance. Relatively unremarked, I've also talked about redescription. Today, I want to start by making redescription considerably more salient than it's been heretofore. So let me start by reminding you of some of the places in Descriptive Psychology where description is a central feature of the situation.

Let me start with our standard significance heuristic. The presentation of this one goes like this: imagine that we're standing outside a farmhouse, a single farmhouse on a lonely heath in Britain. We're standing at some distance, observing, and there's another guy standing fairly close to this farmhouse, and what we see is that he's moving his arm up and down. So if we describe what he's doing, we say he's moving his arm up and down. Now it happens that his hand is grasped around the handle of a pump, so he isn't just moving his arm up and down; he's pumping the pump. Again, as it happens, there's water in the pump, there's water in the well, so he's not merely pumping the pump; he's pumping water. And as it happens, the well is connected with the house, and there are people in the house drinking the water, so he isn't just pumping the pump; he's pumping the water to the people. Further, as it happens, there's poison in the water and he knows it, so he's not merely pumping water to the people; he's poisoning them. Finally, as it happens, those people are conspiring to overthrow the government, and the conspiracy has a good chance of succeeding. So he's not only poisoning those people; he's saving the country.

That's a classic example that appears in the philosophical literature somewhere, and I've forgotten where. But it's noteworthy for the number of different descriptions that you can give of what this man was doing, and that's why I use it: to illustrate that there is no limit, in principle, to the number of descriptions that you can give.

There is something else in the picture, that if we were standing out there watching, probably we would start with one of these two descriptions [moving arm and pumping pump], and from these descriptions, we would move up to these other ones. And it's because there is a characteristic sequence that we talk here about redescription. You describe it here [moving arm]; you redescribe it this way [pumping pump], you redescribe it that way [pumping water to the people] , you redescribe it that way [poisoning the people], and then you redescribe it that way [saving the country].

I want to give you three examples; this is the first one. The second one is old faithful Dinner at 8:30. For those few of you who haven't heard it before, let me go through that one quickly. Imagine that I tell you that last night, I got through work at 6:00 and I got home at 6:30, and we had dinner at 8:30 and it was steak well done. About that time, you yawn and say to yourself, "So what else is new? Half of Boulder could say the same thing." Then I add that yesterday morning I had a big argument with my wife and we never got it settled, and I usually do get home at 6:30 but we usually have dinner at 7:30 not at 8:30, and I like steak, but I like it rare and I hate it well done.

About that time, you have a very different picture of what was going on last night. Initially, the description is simply "We were having dinner, steak well done, at 8:30". After you know those additional facts, you have a new description, namely, she was getting back at me, or something along those lines. So again there is a transition from one description to another description, from the description that we were having dinner at 8:30 to the description that this was a display of hostility.

The third example appears in Decisions and Decision Aids, and it has to do with information gathering. In this example, there are three tanks that move from a given position, A, to another position five miles away, B. That's the description of what happens: three tanks move from A to B. Now as it happens, point f is on a bluff overlooking a fort, so when you take that into account, you have a new description of what happens, namely, the three tanks are threatening that fort. And as it happens, in going from A to B, those tanks cross the national boundary. Taking that into consideration, you have another description of what was going on, namely, there is a general attack going on. Again, you start with a description and then you move to other descriptions.

All of those three are examples of the same general kind, namely, you start with observation and you redescribe it. You give an observation description and then you give a redescription.

Let me switch you to another example. In interpreting Rorschachs, you start with a transcript, a protocol, of what the person said in response to each of the cards, and that's your first description of his behavior: "He said this in response to that". That's not so much observation, because you're only looking at the results. Generally you're not actually seeing him do this; you're simply seeing a record of what he said. It's not a behavioral observation. But you're starting with a description that's easy to give, namely, he said such and such in response to this.

Interpreting Rorschachs requires that you go beyond that description, and redescribe his behavior. For example, you might redescribe it as "He saw something unpleasant", or "He saw something dangerous". Normally, you do at least one more redescription, for example, "He talks like someone who lives in a dangerous world". At some point around that time, you're damn ready to say, "It looks like he's this kind of person": you're ready to go from a behavior description to a person characteristic description, which you then apply and draw conclusions about what he'll do elsewhere and elsewhen. So Rorschach interpretation, in general interpreting projective techniques, requires again the sequence of redescription, starting not from observation but from some product.

Those are the familiar examples. Let me remind you of a couple of other familiar examples that are not so familiar. Think of operationalizing a hypothesis. I want to study hostility, and so I use a five-item questionnaire. My data consists of the responses to the questionnaire, and my first description of the data is exactly that: here it is, this is what they said. The redescription is something about their level of hostility. That kind of redescription is easy because I created that instrument for the sake of being able to do exactly that. I put in those questions or selected those questions precisely so that I could give that redescription, because I'm interested in the redescription. I'm not interested in what they said to those questions; I'm interested in their level of hostility. So in that kind of situation, I deliberately set it up to legitimize a particular redescription that I have in mind. That means that generally, when I do collect the data, I have no hesitation about giving you that redescription, and I will tell you that I found out something about people's hostility.

In contrast, think of the familiar plight of a researcher who collects data and then asks, "What does it mean?" He's looking for a redescription. He needs a redescription. But it isn't the setup any more, and oftentimes redescriptions come hard and you may not find one that satisfies you. But you're looking for that.

Those last two examples should touch a tender nerve, and the nerve is hard data versus interpretation, or in the usual way of talking, hard data versus mere interpretation. So let me make a provoking statement: hard data is absolutely useless. Hard data is absolutely useless by virtue of what makes it hard data, namely, that it implies nothing beyond itself. That's what defines hard data: anything that doesn't imply anything beyond what you already saw, have already established, is hard data. In contrast, anything that's of interest to you, now, is something that has some implications for either now or, more likely, the future. That means that there is a significant problem in moving from hard data to something useful, and how do we do it?

The question "How you do it" arises in some of these other contexts. It certainly arises in the Rorschach class. When people see a few examples of that and pick up the idea that this is what you're supposed to be doing, you can bet that about the first question they have is, "How do you do that? How do you redescribe? How do you move up the significance ladder?" It was partly in reaction against that last time, part of what I said is, "Remember there's no such thing as significance."

What I meant was this: we can talk about moving from here to here to here, or moving from there to here to here, and I do speak of moving up and down the significance ladder. If you said, "Is that what people actually do?", I would have to say Yes and No. But I can put it this way: it's not that you start with this and get this one because you have some way of moving up the significance ladder. Rather, it's because you can generate this one [pumping the pump, from moving the arm] that you can move up the significance ladder.

That's the sense in which there's no such thing as significance. There isn't a procedure for moving up and down that gets you to these places. Instead, it's your ability to generate these various descriptions that achieves the moving up and down on the ladder. But moving up and down on the ladder is not how you do it.

From that you could easily go to the next conclusion, which is that there isn't a way to do it, because it doesn't work that way. It works by being able to do this, and that succeeds, but there's not a way to get from here to there to generate this.

Let me come back to the issue of uselessness. Any data that's going to be useful has some implications for how the world is or might be or will be, somewhere, sometime, not just in the past. That's why something that has no implication beyond itself will have no such application, and that's what makes it thoroughly useless.

In general, how we move from hard data to interpretation is to put it in some kind of human context and say, "In that context, what does this mean for me, or for us?" Notice that this is what we did here [the pump story]: moving up here, we added context each time. It was the additional facts that enabled us to generate the new descriptions. Likewise, in Dinner at 8:30, it was the additional facts that allowed you to say that was an expression of hostility. Likewise, in the case of the three tanks, it was the additional context that allowed you to generate these new descriptions.

The context includes that it's a human context, and you're evaluating those hard facts within the human context, and you're evaluating them in terms of what difference do they make, what importance do they have, what relevance do they have for what I do or don't do? Remember the Judgment Diagram.

What that sensitizes us to, then, is that interpretation is not mere interpretation. Interpretation is what it's all about. Interpretation is where the action is. And interpretation is one of the least developed parts of our research methodology. There's lots of development on the technical, hard-data side, hardly anything beyond folklore on the interpretation side.

I said that you don't literally move up and down the ladder, and that's how come you're able to generate the descriptions. There is another feature that might suggest that, no, nevertheless that is what you do. That is that some redescriptions are easier to generate than others. Take Dinner at 8:30. Given the facts that I gave you, my guess is that 80 to 90 percent of college students get the point and can generate the redescription. That leaves 10 to 20 percent who can't. Think of the gap there is between dinner at 8:30, those three facts, and then saying she was angry at you, she was showing it. It's a gap that most of us will easily bridge, given those three facts.

Now suppose you had somebody who couldn't get from here to there. What could you do to help them reach that redescription? What you can do is give him an intermediate description. For example, if the person you're talking to said, "What do you mean, mad? I didn't see anything that looked like anger." I said, "Look, you know she had reason to be mad at you. She made you wait an hour, and she gave you something she knew you hated. If that's not hostility, what is?" Given that intermediate description—she had reason to be angry at you, she made you wait an hour, she gave you something she knew you hated—given that, it's a lot easier for somebody then to say "Huh, yeah, it looks like she was mad."

It's the fact that there are intermediate descriptions that make the redescription easier, that's part of why it looks like you literally are moving up and down that ladder.

The notion that intermediate descriptions in general make the final description easier to come by is, you might say, one of the crucial principles in interpretation. If you just go from the beginning to the end, like if we saw him moving his arm and say, "Obviously he's saving the country", it's going to look either like black magic, or crystal ball, or like you're making it up, or that you just made a wild guess that happened to be right. As soon as you put in all of the intervening descriptions, there's very little doubt by the time you reach here [saving the country]. It doesn't look like black magic; it doesn't look like you're making it up.

There is another familiar context where you find these mid-level descriptions, and that's in the enterprise known as data reduction. When you calculate means and sigmas, or make up tables or do P-tests, you are generating mid-level descriptions. You're going from the data to some mid-level description that says the means are different—the mean is 5 and this other mean is 4—that's a mid-level description. That makes it easier to see or to conclude, "These two groups are different".

Now for some redescriptions, data reduction is enough. You calculate your means and do your T-tests and do that kind of thing, then it's relatively easy to move from that to the kind of description, the kind of interpretation, that you want to make of it. That's not always possible. So let me introduce explicitly the notion of a redescription structure as a methodological schema. Just to summarize what I've said about that, consider the circumstance where you encounter a phenomenon under a certain description, and that's your initial description. You're interested in that phenomenon under some other description—that's your final description. And it's not obvious how you get from the first one to the second one. Then you introduce an intermediate description which is easier to get to from where you started, and it's easier to get from there to the last description.

If you think of that kind of a situation, where you have the minimum three levels of description, with the function of the middle one being to help you get to the last one, let's baptize that as a redescription structure.

A ——> C via B

Remember that most research fits that structure, because in most research, you do do some data reduction in between.

One of the virtues of considering that in its own right as a methodological schema is that that provides what you might call an exoskeleton that holds together things that don't obviously go together. What I'm going to describe to you is a research project in which exactly that is true, that you can set it up as a redescription structure and that helps pull together things that you wouldn't normally think went together, but in fact you can pull it off.

So let me switch now to an application of this notion. Let me start with some background facts. Part of the background is that astronomy has become almost entirely an electronic subject. In the old days, astronomers used to spend hours looking through telescopes or attaching cameras to telescopes, and that's how you got your primary data in astronomy. Today, very few astronomers spend time either taking pictures through telescopes or looking through telescopes. Instead, they process large data bases with data produced by electronic instruments across a broad range of spectra. You have radio telescopes, you have millimeter telescopes, you have infrared telescopes, you have optical telescopes, you have ultraviolet telescopes, you have X-ray telescopes, you have gamma-ray collectors, and you have cosmic-ray collectors. All of these are processed electronically, so today, becoming an astronomer involves becoming a computer scientist, because all you're ever dealing with are these vast bodies of data, these vast tables of data.

I mentioned—we were talking history, that it was already apparent in 1960 that we were suffering from an information explosion. During the intervening 30 years, roughly 30 years, that information explosion has continued on its way exponentially. We have vastly more data around than we had then, 30 years ago. In astronomy, it's been that in spades. They have vastly more data than they had 30 years ago. In large part, it's been because of the satellites that carry astronomical instruments, and being electronic, they have an order of sensitivity that is much higher than optical telescopes. So being electronic, they can process more data faster; having higher resolution, they can give you more data from the same pointing, from the same observation.

Just to give you an idea of the current dimensions of the problem, there are five warehouses located in the vicinity of Washington, D.C. that are stacked from floor to ceiling with 70,000 magnetic tapes containing data that has come down from the satellites. The person in charge of that whole operation estimated that perhaps 10% of that data will ever be used, because they have no good ways of processing that much data, they have no good ways of making that data accessible to astronomers, and so most of it is just going to stay there in that warehouse until the tapes rot. That's the current conditions.

Now that the shuttle is back in business, there will be six huge telescopes launched at periodic intervals. The first one is the Hubbell Space telescope, and that is primarily designed to take pictures of the planets. It has what is called the wide-field planetary camera on it. Most astronomers are interested in that instrument not because of the planets, but because when it takes pictures of the planets, in the background there are going to be billions of galaxies. The sensitivity of that instrument is orders of magnitude greater than any previous satellite instrument. So it will extend the limits of observation many, many times in terms of distance, and million-fold in terms of how many galaxies this thing can record. The major interest in the instrument, then that it is going to send down observational data on billions of galaxies.

Now the future dimensions of the problem—remember, we're only talking about one satellite and one instrument on this satellite—the current estimates are that that one instrument will send down, in one year, more data than all of the instruments ever used by astronomers. There will be more data than has been gathered in astronomy for the entire history of astronomy, including those five warehouses. And it's only that small because it's throttled down to one-third of its capability because we can't process that much data.

Okay, that suggests the dimensions of the future information explosion problem.

As to the actual data, it is extremely important to a lot of people to be able to classify those galaxies as to their shapes, their morphology. There is a standard classic classification system, and it's called the tuning-fork diagram. The reason is that it looks like a tuning fork.

These are ellipses, and as you move from here (1) to here (4), it becomes a circle. This one (4) is essentially circular, with a nucleus in the middle. From here on out, you have spiral arms. It's no longer a solid thing like these, but all of these have spiral arms. The difference from this one (A) to this one (B) to this one (C) has to do with the spread of the spiral. These (A) are tight and close in; these (B) are a little more dispersed; and these (C) are even more dispersed. Here (D,E, F) it's the same story except that you've got a bar across here. That bar is simply that there's a whole lot of stars there—remember, this is a galaxy. So down at the bottom there are bars on all of these, and in general, the arms extend from the ends of the bars.

Now as I said, it's very important to a lot of people to establish what the shapes of the galaxies are, way out there. There are two major reasons for that. One is, when you get very far out, the shapes of galaxies are about the only clue that people have about what it's like out there. When you get far out, the kind of observation and measurement that you can make close by fuzzes away too much and you can't count on it. What's left is—you can still tell what the shape is, and you can make use of regularities that you've observed in the galaxies close by that relate shape to the local physical conditions. By virtue of those connections, if you know what the shape is of something very far away, you've got a good basis for saying here's what the local physical conditions are, out there.

Naturally, that's of great interest to those astronomers or astrophysicists who are concerned about the whole thing. Remember, Joe talked about now that you can see that far, there are people who are interested in what does the whole physical universe look like. Those are the people who need this kind of information, because that's the only kind they're ever going to get for a long, long time. [laughter]

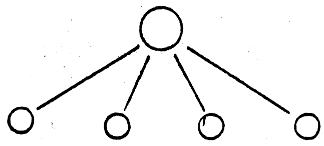
The second reason for wanting to have these things classified is that, as you might expect, not all galaxies fit into these classifications. There are some unusual ones. There are some that don't fit. There are some that are non-standard in various ways. There is a lot of special interest in those non-standard ones. Nobody cares all that much to find another three billion that look like these, but among those three billion are going to be a few hundred or a few thousand that don't fit. Those are going to be the ones of special interest; those are going to be the ones that people will then focus on to study in order to advance our knowledge of what the hell things are like out there. So there are then two crucial bodies of interest in this morphological classification.

The way it's done now, the only way it's done now, is with photographic plates. We've got these nice photographic plates about that big, and there are about half a dozen astronomers in the world, or in the country, to do it. What they do is, they go over with a high-power jeweler's loupe and take a look, and then they have an atlas that has a number of examples like this— paradigm cases, if you will. They look back and forth and they say, "This is one of these", and then they go on to the next. There are in existence catalogues with classifications: one has 16,000 galaxies; another has 250,000. One person has spent her entire career doing nothing but classifying galaxies that way. There are very few people who will do that, and so there are not that many catalogues; there are not that many galaxies that have been classified. The ones that have are all close by. They're the ones that are close by where you can actually see these things, and because you can see them that way, even if they're at an angle somewhat, you can still look at them and say, "This is this kind". That's how it's done.

Now consider the task of classifying those billions of galaxies that are going to be streaming down from that telescope. You look at how it's done now, looking at photographic plates, and you say, No way. You couldn't even make a start. Yet you need to classify a whole lot of them in order to detect the unusual ones. You need to classify the ones furthest away because that's the *terra incognita*. That's what we're most in need of knowledge about. So you're going to have to get beyond (1) doing it by eye, and (2) doing the easy ones. In a word, you're going to have to figure out a way of doing it automatically, and that's easy enough to say, and the question is how.

For our purposes, assume that the electronic data is like a TV screen, namely, it's pixel data. It's data that's organized like a TV screen, namely, some amount of information about this little piece and some amount about the next piece and the next and the next, and so on down the line. That's the data you have to work with. When data comes down from those telescopes, they go through a fairly sizeable operation of simply cleaning up the data by correcting it for the peculiarities of the instrument, the peculiarities of the instrument given the time of day, given the amount of heat on it at the given time—all of these little nitty-gritty things that they know about and they correct for. By the time they get through all of this pre-processing, then you've got clean data that looks like this. That's what you have to work with.. That's your first description. That's your hard data.

hard data



end level

What you want is this kind of description—that's your end-level description. Now you've got a problem, because whatever end you start from, you can't get to the other end.

We know that human judgment can get us here, starting from a photographic plate. It can't get us here starting from pixels. Human judgment doesn't work very well on pixels unless you can see it. But if all you have is a stream of data, human judgment is not very good. It's what you might call non-intuitive. So this is not a task that you can simply appeal to human judgment and human ability to redescribe. Besides, remember it's got to be done too fast. Even if you could do that in principle, it's got to be done faster than that.

The standard ways of dealing with pixel data for this kind of purpose, there's two kinds of things. One is pattern-recognition operations where you simply do analyses of adjacent pixels and see what kind of uniformities there may be. You can generate second-level descriptions just by starting off that way, seeing what kind of shapes you have, what kind of objects. Then you can combine different ones of these in a similar fashion to generate higher-level descriptions. So there are techniques for starting at the bottom level and moving up. However, as you move up, it gets harder and harder: you get diminishing returns because errors multiply, and it becomes harder and harder to write the right rules for this kind of thing. It becomes harder to write the kind of rules that will do the job of the human eye in detecting shapes. It's easy to do the initial stuff because all you've got to do is look at adjacent pixels and decide on certain criteria, and that will give you objects that consist of adjacent pixels. In terms of what Greg Colvin was talking about, you can see that that's the kind of thing a computer can easily do, or should be able to easily do.

In contrast, looking at these things and saying, "Well, that's an FD-2", it's not something that a computer can easily do. It's not something that could be easily programmed to do. There's sort of the boundary conditions of the problem, that you've got the beginning point, you've got the end point, you've got the task of getting from A to B and there's no way to do it. You can start from A but you can't get to B; you can get to B but not starting from A.

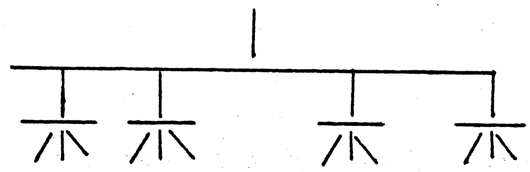
Here's where the notion of a redescription structure pays off. Remember I said, as a schema it will help to hold things together that you wouldn't have thought went together, things that don't naturally go together.

Human Judgment Methods

Syntactic Pattern Recognition

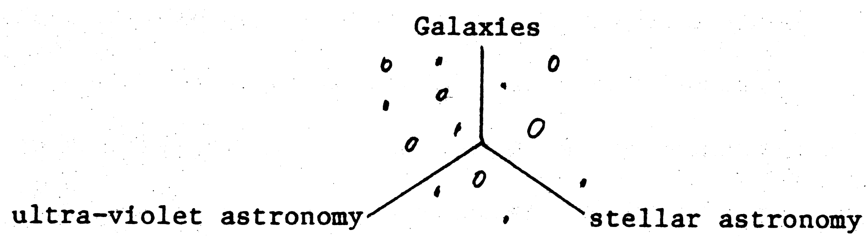
Redescription Structure

Now let me say a little bit about the Human Judgment Method. This is familiar to some of you, but probably not to enough of you to be able to just skip it. Greg referred to this in the context of indexing and retrieving books, or more generally, texts. You think of a classic library setup: it's a branching tree—your subject-matter catalogue in Norlin Library, it has a general description for the whole library, and then it has major subject headings, and under each subject heading it has some number of smaller subject headings, and under each of there is a number of others, and so on up to some point.



When you go look for a book, you have to guess where it is, because if you just went to the library, there's a couple of million volumes, and if you say, "I want something about astrophysics", there's still a quarter million volumes, and if you come down here and say, "I want to know about galaxies", there is a category 'galaxies' down here, but maybe what you want isn't there because it's also about stars and it's been indexed under stellar astronomy. That creates a well-known trade-off, and it's a hopeless trade-off, namely, if you want to be sure that what you want, you have access to, you go up high, like you go up to astrophysics. You can be pretty sure that whatever you want is somewhere in the astrophysics category. However, the higher you go, the more things there are there, and so the more you have to look at, and that's hopeless. You're not going to look through 250,000 volumes. The further down you go, the fewer things you have to look at, but the more likely it is that you'll never find what you want because it's somewhere else.

The solution to that problem was, instead of grouping things under categories, you spread them out along dimensions. You could have exactly the same dimensions as you have categories, for example, galaxies, stellar astronomy, ultra-violet astronomy, and for each document, instead of saying, Which pigeon-hole does it belong in, you get some person who knows about these things to say, "This one is to this extent relevant to galaxies. To this other extent, it's relevant to stellar astronomy, and to this extent it's relevant to ultra-violet astronomy." What you wind up with, as a result, is that you've got a set of numbers that you can transform into co-ordinates in a space. You've got a spatial representation, and each document has a location. And as Greg showed earlier, the way you get documents back, you make the request; it gets indexed in there, and whatever is close by is likely to be what you want.



You can come back with these things in order of distance, and that makes it easy because you now no longer have 250,000; you simply have the closest and the second most close and so on. You just read down the list, and the chances of your finding out what you want are pretty good, in practice.

To re-index documents, you can do it by having a human read them, and then rate them. However, that's very inefficient, and so there is an additional procedure for doing it automatically. That procedure is to get yourself a moderately large set of terms that you have reason to believe will appear in these documents. Each of those terms, you get the same kind of judgments on, so that each term has a location. Then when you get a document, you simply scan it, determine which of these terms appear in it. Since each of the terms that appear in it have location, you then give the document a location based on the locations of the terms that appear in it. That way, you can just run a stream of texts through and automatically index from now until kingdom come, and it all happens untouched by human hands. Once you get the original human judgments, the thing will operate completely automatically.

What do these things represent—what do these point locations represent? They represent a subject-matter profile. You think of this not as a spatial location; the single point can be thought of as a profile, and a profile of subject-matter relevance.

Now a little-known fact about this arrangement is that indexing in this space is a redescription in the very sense that I've talked about. What you encounter when you're processing the document is not a word. What you encounter is a character string. You can recognize the character string; you cannot recognize the word. What these judgments do is to give you a redescription of this character string as a word that has this subject-matter profile. Now you can deal with it as a word that has this subject-matter profile; you don't have to deal with it as a character string any more. This kind of arrangement, then, is a way of automatically giving you that kind of redescription, and its particular virtue is that it can give you the kind of redescription that only people can do.

Q: Would you go over that one more time ——?

Peter: As a word that has a subject-matter profile. Then you can deal with it as a word that has this subject-matter profile rather than having to deal with it as a character string. Keep in mind that the standard way of retrieving is to recognize the character string, and then look in your documents, or have a record of which other documents have the same character string in them. And the character string does not have to be a word. It can be any character string. That's the principle on which key-word retrieval is done. The way that this is different is that you get beyond the character string and work with it as a subject-matter profile.

At this point, you have half the answer. You could hope to get this kind of description at the top level, using that kind of human judgment system to get you there from a lower level. Because that's what that does: as soon as you index, you've bounced up one level. You've automatically generated a redescription. So you can set up a system where your dimensions are not these [tuning fork] but rather these things [Galaxies, Ultra-violet astronomy, stellar astronomy diagram], or their parameters. And in fact, it's better if you don't use literally these, but rather have their parameters so that you've got a space, and this kind of thing routinely appears up here, this kind of thing appears somewhere else, and this one appears somewhere else, so that by where it is, you're going to be able to tell which of these it is. The reason why this is preferable is that here, you're going to find a group of things that aren't any of these but there's clearly a group of them, and that's how you're going to recognize your unusual cases that are going to repay lots of further study.

With that kind of methodology, you can hope to arrive at this top level from some lower level. And now you can see where it's going, that if you start with pattern recognition at the pixel level and move your way up to some intervening description, and if you start at the top with this kind of system and work your way down by asking What does it take to recognize which of these it is, and you get a number of different descriptions, if it's this and this and this, then it's one of these. Then, just like Joe's diagram comes down, you say, If you need this and this to recognize that it's one of these, how do you find out if it's one of these? How do you find out if it's one of these? The answer is simple: if you just put these in, into another one of these where your redescription is going to give you these, starting from a still lower level of description. At some point, the kind of lower-level description that you're working with is exactly the kind of thing that you're generating from the bottom up by pattern recognition. Somewhere those two things are going to cross enough so that you can then start from here [galaxy diagram] and end up here [tuning fork].

Q.——how is it working?

Peter. It's not built. It's still in the planning stage. That's the in-principle solution. Here's how you can go about it, and automatically classify those things as to the strings come down from the telescopes, and wind up with a nice, neat separation of the ones that you know about, and hope to discover the ones that you don't know about.

After the fact, it looks easy. It looks almost obvious to do that thing in steps, partly, though, because we're used to going through steps of redescription. From the standpoint of contributing to the science, this is going to do a whole lot for them if it works. It's something that has not been tried before; it's not part of the standard methodology of astronomy or astrophysics, but it may become. If the system works, it damn well will become part of the standard methodology of astrophysics, because there's a whole lot of other places where you can do similar things.

In a sense, we could stop there, but let me tell you that it's even harder than this. It's easy to do it when you've got that much information, when in effect you're looking at them broadside and it's close enough to be big enough that you can see these things. Imagine, just like that mystery picture, that you systematically degrade this, the further and further the thing gets, the less you're going to be able to see detail, the more it's going to look like that, and finally it's just going to sort of look like this because you're going to push this telescope to the limits of its resolution.

Part of the problem is to be able to deal with these and not just these, and to allow yourself as much error as you want, but you require that on the average, if it's a spiral like this and it's far away, on the average you call it a spiral like this. You can be wrong a lot of times, but as long as on the average you give it this, that's okay because then the whole thing is consistent.

Now if when you get down to very degraded cases, that's going to put a lot of strain on the technique that had been developed for the easy cases. So there is lots of additional technical problems that we can foresee, but every time we look at the in-principle solutions, they look better and better to us. It looks more and more do-able as we get down to more detail and say, How would you actually do it? What dimensions would you use?—enough of that to say, Yeah, boy, it sure looks do-able.

This won't be the first time that psychologists can help the astronomers. Remember the original reaction-time experiment. You might say this is an interesting methodological exercise; it's also an interesting excursion into another science and their problems, which turn out to be not that different from our problems.

I think that's about as much as I want to say. Any questions?

Q. [mostly unintelligible, something to do with musicology and recognizing melodies]

Peter. As soon as you get the general principle, you can just think of all kinds of possible applications for it.

Q. I heard you say "we" ——-

Peter. I didn't think this up by myself. I have a collaborator, somebody from the Harvard-Smithsonian Observatory, and I spent a whole day talking with him about the measurement problems of astronomy, and he kept coming back to this one as one of the most important. So cyclically, we came back and said, What could you do about that? And after a few tries, we came up with it.

Q. Where does this stand in terms of ——-

Peter. Two things. One, we have another project that has to take priority and that will occupy us at least until the end of the year and maybe further. Secondly, this thing is going to cost, and we need to get funding.

Okay. [applause]