

CUSTOM

Y-UCC
DD

AUTHOR

Market for a Generalized Data
Dictionary

INPUT

PAC
copy - 1

701 WELCH ROAD, SUITE 1119, PALO ALTO, CALIFORNIA, 94304 (415) 854-3422

FINAL REPORT

ANALYSIS OF THE MARKET
FOR A GENERALIZED DATA DICTIONARY

for

University Computing Corporation

on

December 8, 1975

prepared by

INPUT
701 Welch Road
Suite 1119
Palo Alto
California 94304

INPUT LIBRARY

TABLE OF CONTENTS

		<u>Page</u>
I	Introduction	1
	Method of Research	1
	Characteristics of Respondents	4
II	Summary and Recommendations	11
	Recommendations UCC 10	11
	Recommendations Generalized Data Dictionary	12
	Other Recommendations	13
	Summary of Key Findings	14
III	Market Analysis	16
	DP Environment of Data Dictionary Users	16
	Market for a Generalized DD	38
IV	Technical Requirements	44
	Analysis of Data Dictionary User Response	44
V	Marketing Requirements	66
	Justification for, and Value of, a Data Dictionary	66
	Pricing	77
	Support Requirements	77
VI	Competition	83
	Respondents' Concerns with Data Dictionary Vendors	83



Digitized by the Internet Archive
in 2015

https://archive.org/details/analysisofmarketunse_2

LIST OF EXHIBITS

<u>Exhibit</u>	<u>Title</u>	<u>Page</u>
I-1	List of Interviews	2
I-2	Respondents by Industry Sector	5
I-3	DBMS and DDs Used by Respondents	6
I-4	Respondents by size of Installation	8
I-5	Respondents by Operating System	9
III-1	DD Use by Major Programming Language	17
III-2	On-Line Systems Software Use	19
III-3	DD Use by Programming Staff Organization	21
III-4	DD Use Related to Standards Definition Function	22
III-5	Respondents with Standardized/Centralized File Definition and Maintenance	23
III-6	Respondents DD use by Programming Effort in Maintenance	25
III-7	Respondents' Use of DBMS and DD	26
III-8	Use Of DD by Length of Use of DBMS	28
III-9	Attitude of Respondents to DD	30
III-10	Large Respondents DD Use and Intentions	33
III-11	Analysis of Large Users who will get a DD	35
III-12	Very Large Respondents DD Use and Intention	36
III-13	Analysis of Very Large Users who will get a DD	37
III-14	Relationship of Centralized and Specialized DDs	39
III-15	Probability of Respondents buying a Centralized DD	40
III-16	Forecast of IBM sites' Changes 1976/1979	42
III-17	Potential Penetration by Generalized Data Dictionary	43
IV-1	Required Features of a Data Dictionary	45
IV-2	UCC 10 Feature Analysis	47
IV-3	IBM DD Feature Analysis	48
IV-4	Features Analysis for users with their own DD	49
IV-5	Other DD Users Features Analysis	50
IV-6	Respondents evaluation of Utility Features/Interfaces	51
IV-7	Respondents Evaluation of Documentation and Reporting Features	53
IV-8	Respondents Evaluation of Programming/ support Features	54
IV-9	Respondents Evaluation of Operational Characteristics	55

TABLE OF CONTENTS (continued)

<u>Exhibit</u>	<u>Title</u>	<u>Page</u>
IV-10	How Respondents Enforce Programmer Use of a DD	57
IV-11	Respondents required Interfaces with Software Packages	58
IV-12	Operational Problems Encountered by DD user	60
IV-13	Use of DD for all or parts of Respondents Data	61
IV-14	UCC 10 - Generalized Data Dictionary	65
V-1	Justification Criteria for a DD Advanced by Respondents	67
V-2	Respondents evaluation of selected criteria	68
V-3	Relationship of a DD to a DBMS from Installation Date Point of View	69
V-4	Respondents Evaluation of the Real Value	71
V-5	Time Taken for a DD to be of Real Value	73
V-6	Respondent Evaluation of a Package Selection criteria for a DD	75
V-7	Selection and Purchasing Authority for a DD	76
V-8	Expected Price of a DD Package	78
V-9	Suggested Pricing Chart	79
V-10	Respondents Need for DD Source Code	82
VI-1	DDs Considered by Respondents	84
VI-2	Respondents Comments on Vendors & DDs	85

I. Introduction

Method of Research

o This study of the market for a generalized data dictionary (DD) was performed by INPUT during November 1975.

o An initial meeting was held at UCC in Dallas on October 28, to define the questions to be addressed in the study and the interview procedures. An interim meeting was held on November 20, at which some of the preliminary results were discussed.

o The user research was carried out as follows:

On-site interviews	6
Telephone interviews	44

Exhibit I-1 presents a list of the companies included in the interview data base. One company (Fireman's Fund) was independently interviewed twice.

o Other interviews were held with vendors, and with small computer users to determine their knowledge of data dictionaries.

o The definition of a data dictionary used for the study was:

A computer system for standardizing, controlling and maintaining definitions and descriptions of data elements.

A generalized data dictionary is one that is independent of any data base management system or language.

o The market emphasized in the research was that for an IBM-based product; emphasis was further placed on OS installations rather than DOS.

o UCC actively participated in the search for respondents and provided data on a number of them through 'Research Forms' completed by the sales force. These leads were handled as follows:

Research Forms Submitted	42
Interviews	9
Contacts	7
No Contacts	26

EXHIBIT I-1
LIST OF INTERVIEWS

Number	Respondent	Staff	Industry
01	City of Long Beach	HAS	State/Local
02	Pacific Mutual	HAS	Insurance
03	Union Oil Co. of California	HAS	Process Mfg.
04	Rockwell, Computing Services	HAS	Discrete Mfg.
05	United Gas Pipeline Co.	HAS	Process Mfg.
07	Combined Ins. Co. of America	HAS	Insurance
12	University of Minnesota	HAS	Education
13*	Hewlett Packard	PAC	Discrete Mfg.
14	Sandia Labs	HAS	Other
15	Norton Co.	HAS	Discrete Mfg.
16	Indiana NYC Bank	HAS	Banking
17	New York Telephone	PAC	Utilities
19*	Bank of America	HAS	Banking
20	Burlington Northern	TXC	Transportation /Rail
21	Fireman's Fund	HAS	Insurance
23	Atlantic Richfield	HAS	Process Mfg.
25	Miller Brewing Co.	HAS	Process Mfg.
26	Employers Ins. of Wasau	HAS	Insurance
27	Northwestern Mutual Life	HAS	Insurance
28	GTE Sylvania	HAS	Discrete Mfg.
29	Spargue Electric	HAS	Discrete Mfg.
30	Brown Group	HAS	Discrete Mfg.
31	Confidential	HAS	Discrete Mfg.
32*	Wells Fargo	PAC	Banking

EXHIBIT I-1 Continued....

Number	Respondent	Staff	Industry
33*	Utah International	PAC	Other/mining
34	Transamerica Inform. Services	PAC	Insurance
35	J.P. Stevens Co.	PAC	Discrete Mfg.
36	John Deere & Co.	PAC	Discrete Mfg.
37	County of Contra Costa	PAC	State/Local
38	Crown Zellerbach	PAC	Process Mfg.
39	ERDA/NBS	PAC	Federal Govt.
40	PGE	PAC	Utilities
41	La Centrale	PAC	Banking
42	Accacia Mutual Life	TXC	Insurance
43*	Ashland Oil Co.	TXC	Process Mfg.
44	E.I. Du Pont de Nemours	TXC	Process Mfg.
45*	Equitable Trust Co.	TXC	Banking
46	First Computer Services Inc.	TXC	Banking
47	Giant Food Inc.	TXC	Retail
48	Hercules Inc.	TXC	Process Mfg.
49	Franklin Mint	TXC	Process Mfg.
50	Washington DPA	JJS	State/Local
51	Fireman's Fund (2nd interview)	JJS	Insurance
52	Kings County	JJS	State/Local
53	Washington Gas & Light	JJS	Utilities
54	Riverside County	JJS	State/Local
55	University of Washington	JJS	Education
59	Humana	PAC	Medical
65	Varian	JJS	Discrete Mfg.
68	GEICO	TXC	Insurance

Other Leads Given		35
Interviews	21	
Contacts	7	
No Contacts	7	
Total		77
Interviews	30	
Contacts	14	
No Contacts	33	

o Interviews included in the data base were carried out by the following people:

Tony Constable	10
Peter Cunningham	13
Herb Seidman	20
John Shea	7

Characteristics of Respondents

o Exhibit I-2 shows the interview distribution by industry and those respondents using a DD.

o The sample is so skewed by other factors that it is not possible to derive any industry-related propensity to use a DD from it. However, the search for respondents indicates that the following industries may be considered higher potential than the average:

Manufacturing
Utilities
Banking
Insurance

Low potential may be expected in:

Retail
Wholesale
Government

o For a variety of reasons the survey covered more IMS users than had been planned as shown in Exhibit I-3. Some of these reasons are:

- Interviews were primarily held with organizations with very large IBM installations - the penetration of IMS is very high in these organizations.
- Computer users that had not been expected to have a DBMS turned out to have acquired IMS or another DBMS recently. However, 22 of the 26 IMS users were listed on the UCC computer print out.

EXHIBIT I-2

RESPONDENTS BY INDUSTRY SECTOR

<u>Industry Sector</u>	<u># of Respondents</u>	<u># Using DD*</u>
Process Manufacturing	8	7
Discrete Manufacturing	10	9
Transportation	1	1
Utilities	3	3
Banking	6	4
Insurance**	8	5
Medical	1	-
Education	2	1
Retail	1	-
Wholesale	-	-
Federal Government	1	-
State/Local Government	5	1
Services	-	-
Other	3	2
	49	33

* Included in the 'using' figures are respondents developing their own DD or with firm orders for a DD.

** Two interviews were held independently with one respondent.

EXHIBIT I-3

DBMS AND DDS USED BY RESPONDENTS

Total Number DDS	26	9	3	2	4	9	53
Non-User	5	1	1	2	1	6	16
Pride-Logic		1					1
Lexicon	1					2	3
Data Catalog	2	2			1		5
Own	3	4	1		2	1	11
IBM	8		1*				9
UCC 10	7	1*					8
	IMS	TOTAL	ADABAS	IDMS	OTHER	NON- USERS	TOTAL NUMBER

DATA BASE MANAGEMENT SYSTEM

* Respondent also has IMS

- In order to cover the DDs and people that were knowledgeable on them, known DD users were selected for interview; many of these were IMS installations.

- o Four respondents had two DBMS: they had:

- IMS/TOTAL (2)
 - IMS/ADABAS
 - IMS/User Developed

In both of the first two cases, TOTAL was being phased out and replaced by IMS.

- o A major surprise was the number of respondents that had their own DD. Of the 11 identified, two were in the process of developing their own DD and a third had IBM's DD as well as one of their own on a DEC 11/70.

- o An analysis of the respondents by size of computer in the installations is shown in Exhibit I-4. Respondents were classified by the largest mainframe in the installation as follows:

- Very Large - IBM System /370 Model 158 or larger
 - Large - IBM System /370 Model 145 and 155,
and System /360 Models 50 and 65
 - Medium &
Small - Less than IBM System /370 Model 145

- o Two respondents with non-IBM hardware actually had:

- Burroughs B6700
 - CDC 6600/Univac 1108

One other respondent, a university, had a System /370 Model 145 as its largest IBM computer, but had other large-scale CDC equipment.

- o In Exhibit I-5, the analysis of respondents' installations by operating system is shown. Of the very large installations only one of the OS/MVT and MFT users is not planning to upgrade to VS. Also, three of the existing respondents are already on MVS while 10 plan to move to it.

- o There is less activity with the large users of whom only three are planning operating system upgrades to VS. Two others are considering it while a remaining installation is going to fully implement VM which is now under partial use.

EXHIBIT I-4

RESPONDENTS BY SIZE OF INSTALLATION

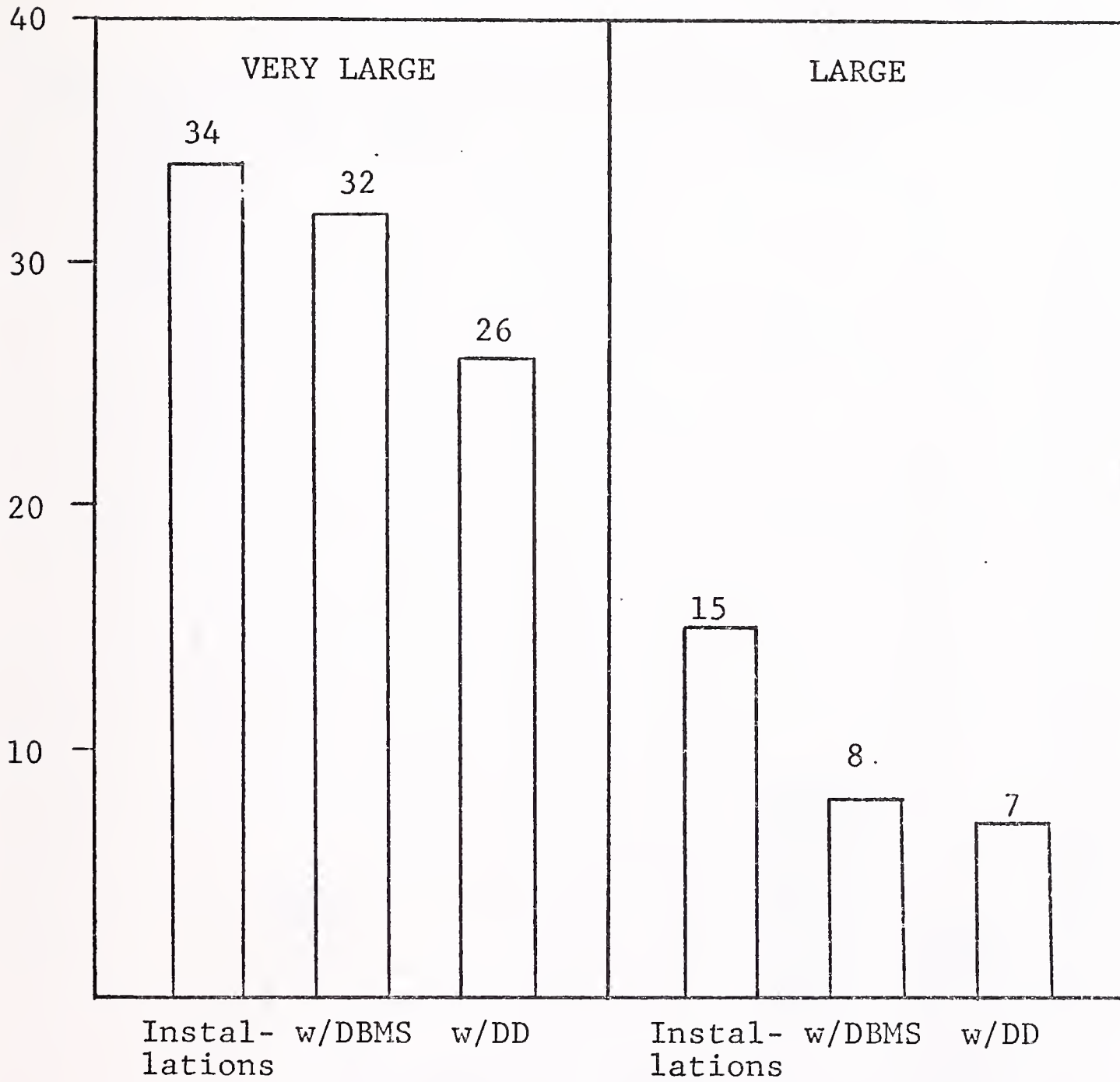
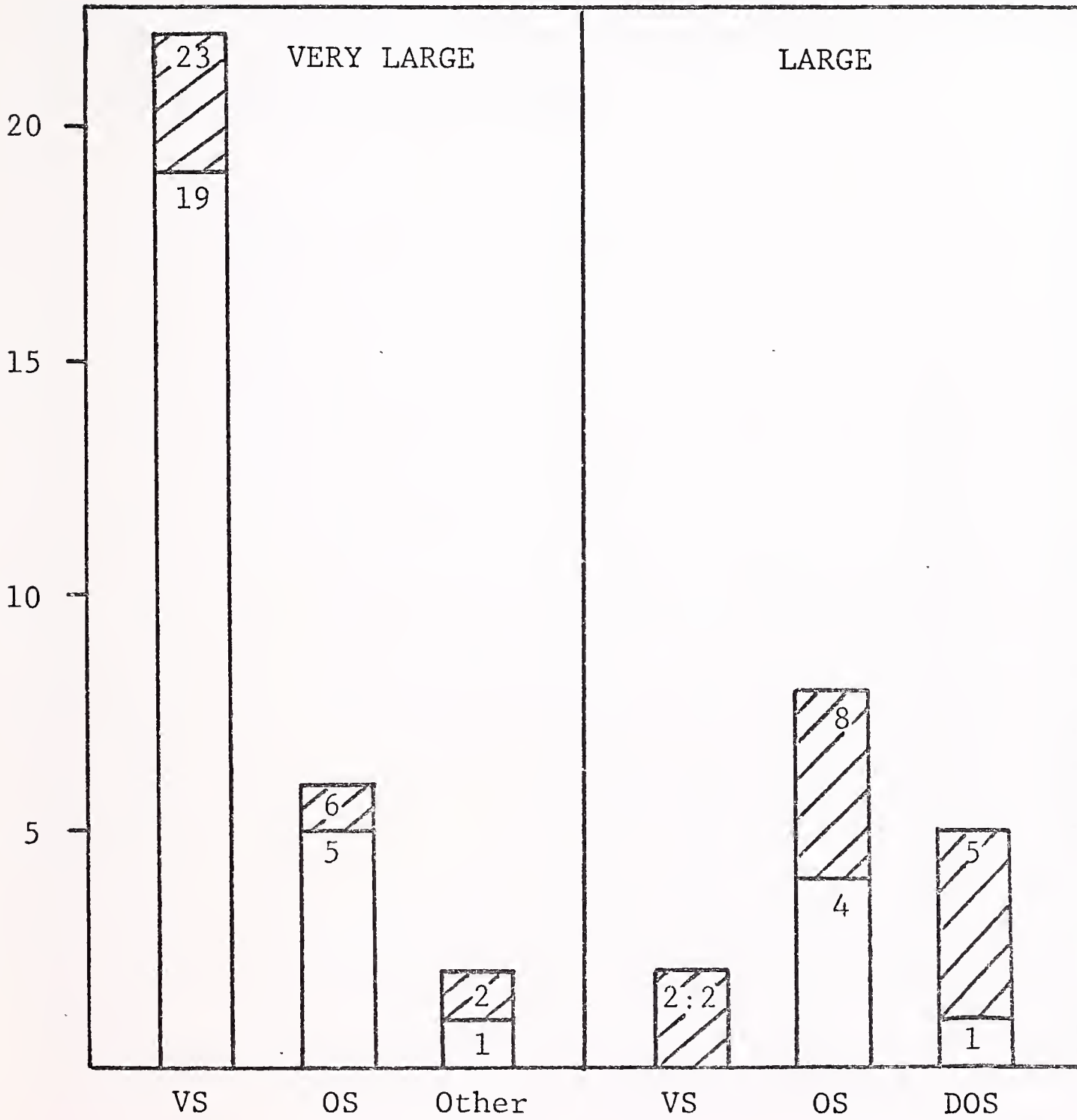
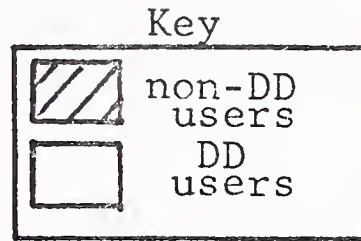


EXHIBIT I-5
 RESPONDENTS BY OPERATING SYSTEM



o From the information collected there is little evidence of any trend in variations in use of DD by operating system within equipment size, except that only one out of five DOS users in large installations had a DD while six out of 10 OS or VS users had one.

o From the analysis of operating systems, it is apparent that the respondents are, on the average, ahead of the bulk of installations in their level of sophistication. They can, therefore, be regarded as 'leading edge' to a limited extent.

II SUMMARY AND RECOMMENDATIONS

Recommendations - UCC 10

- o UCC 10 as currently constituted is a superior product to its main competitor (IBM) in the IMS-based data dictionaries. However certain improvements will improve its position vis-a-vis IBM still further and also impinge on the market for a more generalized data dictionary.
- o INPUT recommends that UCC upgrades UCC 10 as follows:
 - Improve documentation - make it more useable particularly for non-DP oriented managers
 - Improve reports and report generation; again reports should be useable by non-DP people
 - Remove UCC10 from the IMS operating environment so that users do not have to put IMS up before they can use UCC 10
 - Provide interfaces with standard report writers such as CULPRIT, EASYTRIEVE, and ASSIST, and also with GIS
 - Provide interface with PANVALET
 - Provide interfaces with IMS design aids such as DB. PROTOTYPE and DB MAP
 - Provide table processors for non-DP data
 - Provide for source code input from PL/1 and COBOL programs
 - Improve input procedures and forms to simplify use
- o Marketing of UCC 10 needs emphasis in the 'futures' area. IBM is currently outselling UCC 10 by promising features and support in the future as well as 'integration' into IMS.

INPUT

- o INPUT recommends that UCC emphasize that: IBM will keep the DD as an option and that there will be a fee charged for it which is at least as high as it is now; UCC will develop UCC 10 to parallel IMS developments in file structures and facilities; it will be as integrated as IBM's option; and furthermore UCC 10 will provide interfaces and features which IBM cannot do, such as PANVALET.

Recommendations - Generalized Data Dictionary

- o Users want the specialized features of a DBMS related DD but they also want more generalized features. Given the choice of a specialized or generalized data dictionary almost all IMS users would choose the specialized one - they need it more. However, a generalized DD with IMS features will achieve an increasing penetration in IMS-users as they try to cover more with their DD. Non-IMS users will probably choose the generalized one, provided it has features to relate to its (non-IMS) DBMS of choice.

- o Therefore INPUT recommends that UCC should continue to emphasize UCC 10 for its IMS users and obtain a generalized DD for non-IMS users or IMS users which want to cover non-IMS data and systems.

- o The prime generalized DD is Data Catalog; although missing a number of features, including automatic relationships with IMS and TOTAL it has enough capability that some IMS users are selecting it over IBM and UCC10. With certain significant improvements it should be very hard to beat, as a generalized DD.

INPUT's recommendation is that obtaining Synergetics or rights to the package would be a preferred course of action.

- o DATA MANAGER has some good features but appears to be more limited in its relations with DBMS systems.

INPUT recommends that DATA MANAGER be regarded as the next source of a generalized DD. It needs to be able to interface with IMS and TOTAL in an automated way before it can be regarded as a competitor to DATA CATALOG.

- o An alternative approach is to generalize UCC 10. Given that it is supposedly cumbersome and totally oriented to IMS, it is still well regarded as far as features are concerned by all non-IMS respondents we talked to. Removing it from the IMS operating environment will accomplish a major step in generalization.

Also, most respondents with a specialized DD will consider expansion of it to cover generalized data. Thus the generalization of UCC 10 will be well received by existing users; they would probably pay for the new version. It will also be a competitive thrust against IBM.

Therefore, INPUT recommends that UCC 10 be 'generalized' and, if other courses of action fail, used as the basis for a generalized DD.

- o One UCC 10 user was concerned that UCC could fragment its support by introducing another DD. Also, in the long run, there will be only one market for DDs, rather than separate generalized and specialized ones. The degree of specialization required is greater for IMS than other file management systems yet all generalized DDs will support IMS eventually.

Therefore INPUT recommends that the long term objective of UCC should be to have one DD product consisting of several modes and modules.

Other Recommendations

- o Data Dictionaries should be marketed from the 'top down'.
- o Prime targets have some or all of the following characteristics:

- Have had a DBMS for a year or more

- Use PL 1
 - Are engaging on a major development effort
 - Upgrading equipment/operating systems, particularly from a 370 Model 145 or 155 to VS on a Model 158 or greater
 - Implementing a standards function
- o UCC should prepare and use an executive-level presentation on data dictionaries and what they do. This would be an excellent support tool for UCC 10 and other DD clients. It would also be an important sales tool. This is because a DD represents a different way of doing their business.
 - o Pricing of DDs should be on a modular basis. This would allow a first-time user to get in more cheaply, without all the 'bells and whistles'; whereas the sophisticated user can get the options it wants. Such modules would include DBMS interfaces, on-line module, and project management capability.
 - o The major sales point for DDs should be 'control'
 - o UCC should immediately investigate the potential use of DDs in relation to regulation of privacy, security, and the opening of government data to public inspection. This may well provide a major new market.

Summary of Key Findings

- o Data Dictionaries will be a standard feature in almost all very large installations by 1978.
- o DD use is inextricably linked with DBMS: the overwhelming majority of DBMS users consider a DD should be implemented before a DBMS.
- o Of very large companies, as many as 25% may have their own DDs; a further 25% are likely to have some rudimentary file definition

manager. This represents a confirmation of the need for DDs, and their market potential.

- o Users of DBMS other than IMS are less likely to require the degree of specialization in DD that IMS does. Thus generalized DDs are more appropriate products to these DBMS users, except where the DD is tied in with the DBMS sale.

- o IBM is the main competitor to UCC 10 in the IMS DD market: it is not as 'good' a product but IBM is selling 'futures'.

- o DATA CATALOG is achieving sufficient IMS capability that IMS users are now considering it with UCC 10 and IBM. It is the standard against which any generalized DD must be measured.

- o Arthur Anderson is doing a nice marketing job with Lexicon. It also will have IMS and TOTAL capabilities. Its value as a competitor is reduced by the way it is sold: it is a 'free' product which accompanies a significant (\$50,000) consulting contract.

III MARKET ANALYSIS

DP Environment of Data Dictionary Users

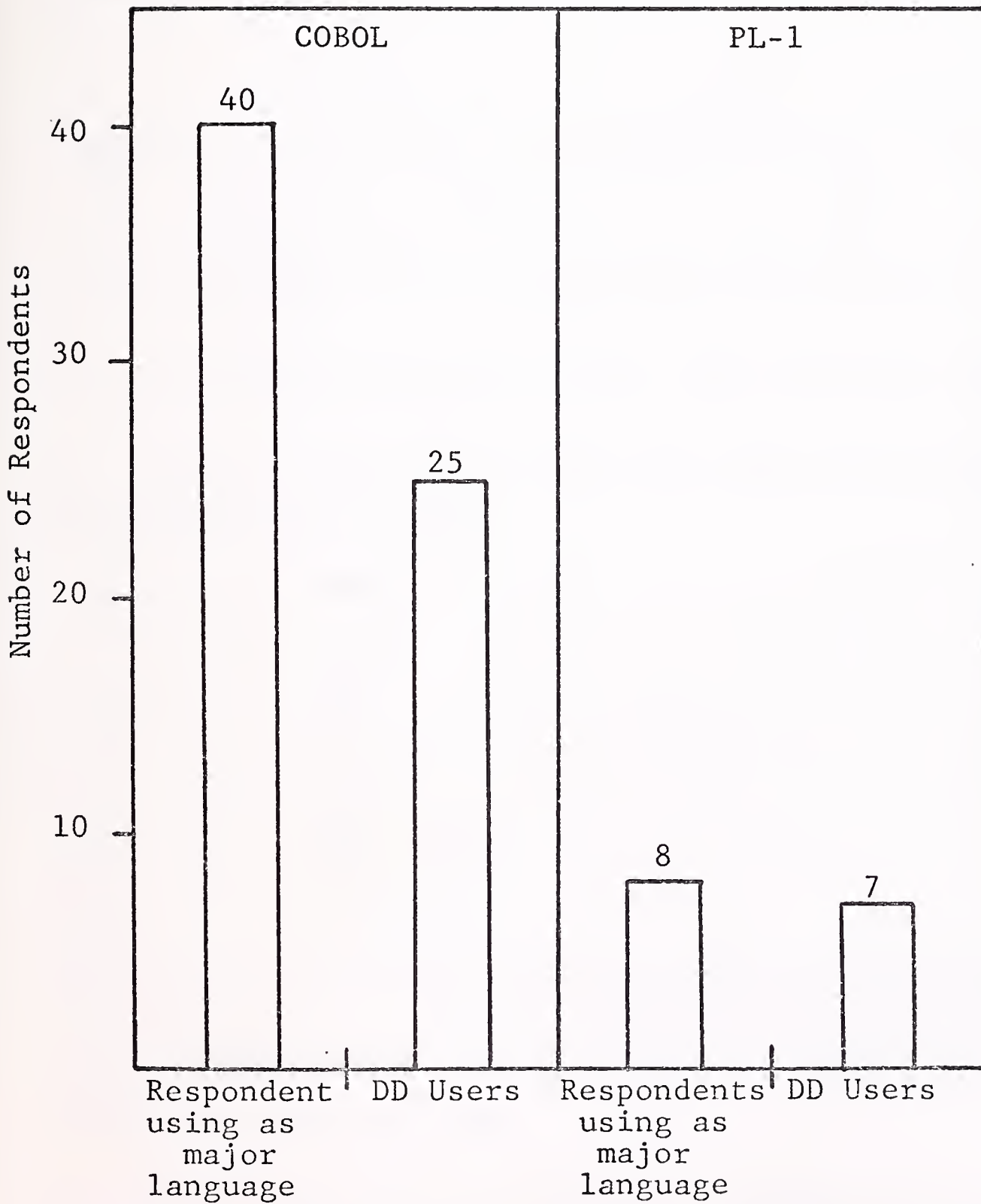
- o Of the very large and large computer users interviewed, all of those who responded were familiar with the concept of a DD. This was not unduly surprising because of the target selection.
- o None of the medium sized installations contacted were familiar with the concept.
- o All of the very large and large respondents would have an interest in such a product: 12 of the 30 respondents to the question 'would you have an interest in such a product (a DD)?' do not now have one on order or propose to order one.
- o Of the six respondents which were now considering data dictionaries and had:
 - selected one,
 - ordered one, or
 - were developing one of their own,

five of these had significant activity in their equipment and/or operating system area. Three of these were replacing or had just replaced a 370/145 with a 370/158, another was adding a 370/158, and a fifth was upgrading its operating system while doubling the memory size on 7/5 158s. Only one of the six intended to retain its present configuration.

Thus, there seems to be a strong relationship between major equipment/operating system upgrades and an interest in DDs, particularly for these

EXHIBIT III-1

DD USE BY MAJOR PROGRAMMING LANGUAGE



companies crossing the larger (370/145 and 155) to very large (370/158 and up) user dividing line. Although, not identified in the survey, this upgrade and the accompanying interest in a DD is probably due to large system development efforts.

o As shown in Exhibit III - 1, there appears to be some relationship between the use of PL1 as a programming language and a propensity to use a DD. Of the eight respondents using PL1 as a major language only one did not now have a DD and it was in the process of acquiring one.

o Of four other minor users of PL1, only one was not using a DD.

o Of the 12 respondents who used PL1 to any extent, 10 of them had a DD, distributed as follows:

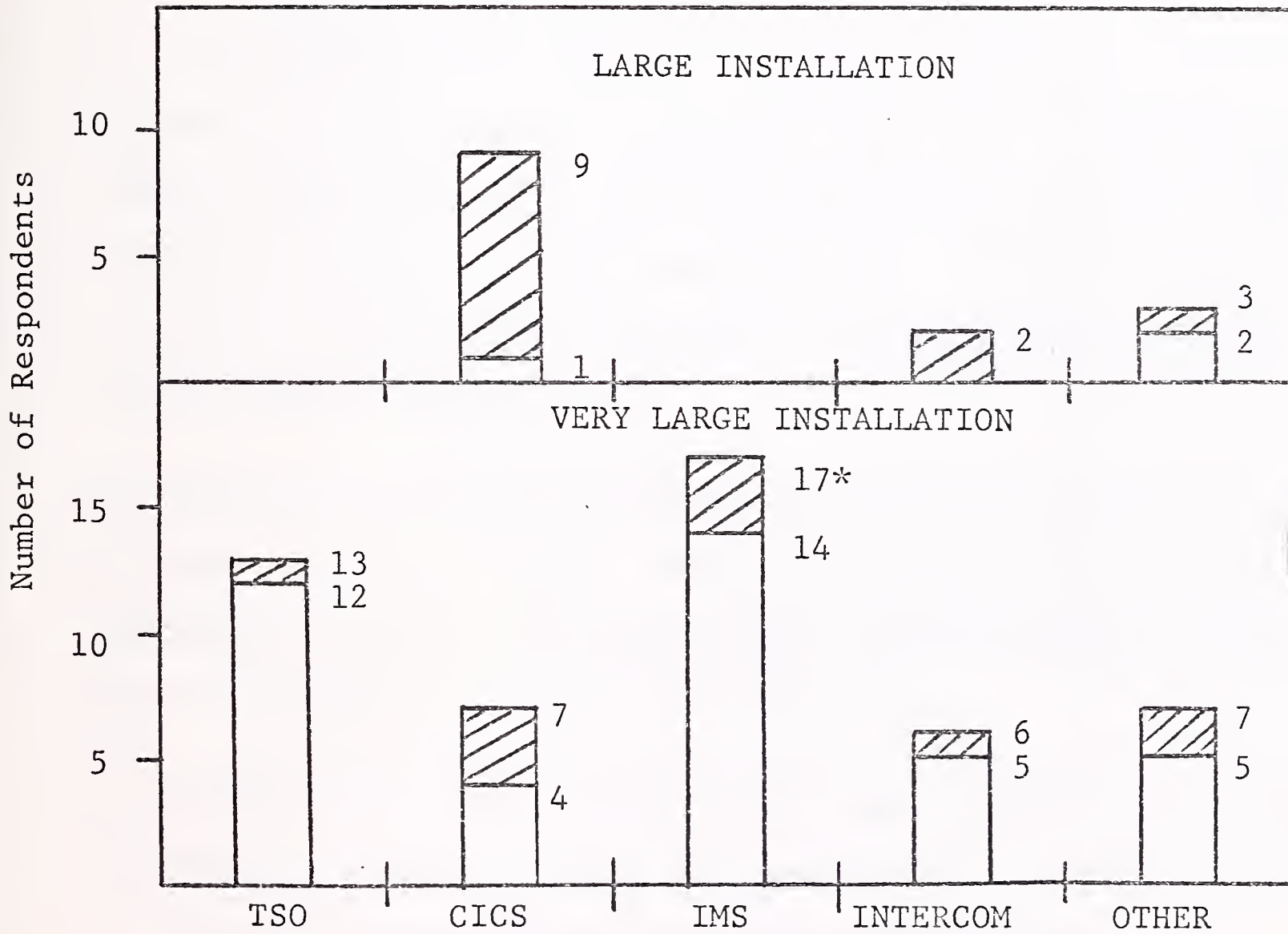
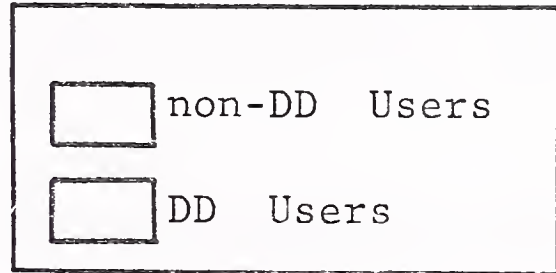
IBM	3
UCCIO	2
OWN	2
LEXICON	1
DATA CATALOG	1
PRIDE LOGIC	1

o As shown in III - 2, CICS is by far the most frequently used on-line system software for large users. However, of 16 large and very large CICS users, only 5 had a DD.

o On the other hand, all or almost all INTERCOM, TSO and IMS users were using or planned to use a DD.

EXHIBIT III-2

ON-LINE SYSTEMS SOFTWARE USE



*All three have DD proposed or in process.

- o Since very large organizations often have decentralized programming staffs, it would be expected that a high percentage of respondents would fall into this category. In fact, 35% of the very large respondents had decentralized programming staffs, and of these all but one had a DD, as shown in Exhibit III - 3.
- o Only one large respondent, LaCentrale, had a decentralized programming staff; it, of course, uses UCCIO.
- o A surprisingly small number of respondents reported an in-place standards definition function, as shown in Exhibit III - 4. The responses indicate, as may be expected, some relationship between the use of a DD and the existence of a standards definition group. However, it is certainly not true to say that DD users always have a strong standards definition function, based on this research.
- o At least six of the respondents were attempting to upgrade or install a standards definition function. Three of these were in the process of acquiring a DD. Thus, there appears to be a strong relationship between increasing emphasis on standards and getting a DD.
- o Reinforcing the concept of the DD as a weapon in the 'war' of standards implementation were the responses to the question:

"Have you standardized/centralized file definition and maintenance? (Did you before getting DD?)"

As shown in Exhibit III - 5, 11 respondents felt that this function was or will be primarily fulfilled through the use of a DD. In the words of one respondent, it is the 'hammer'.

EXHIBIT III-3

DD USE BY PROGRAMMING STAFF ORGANIZATION

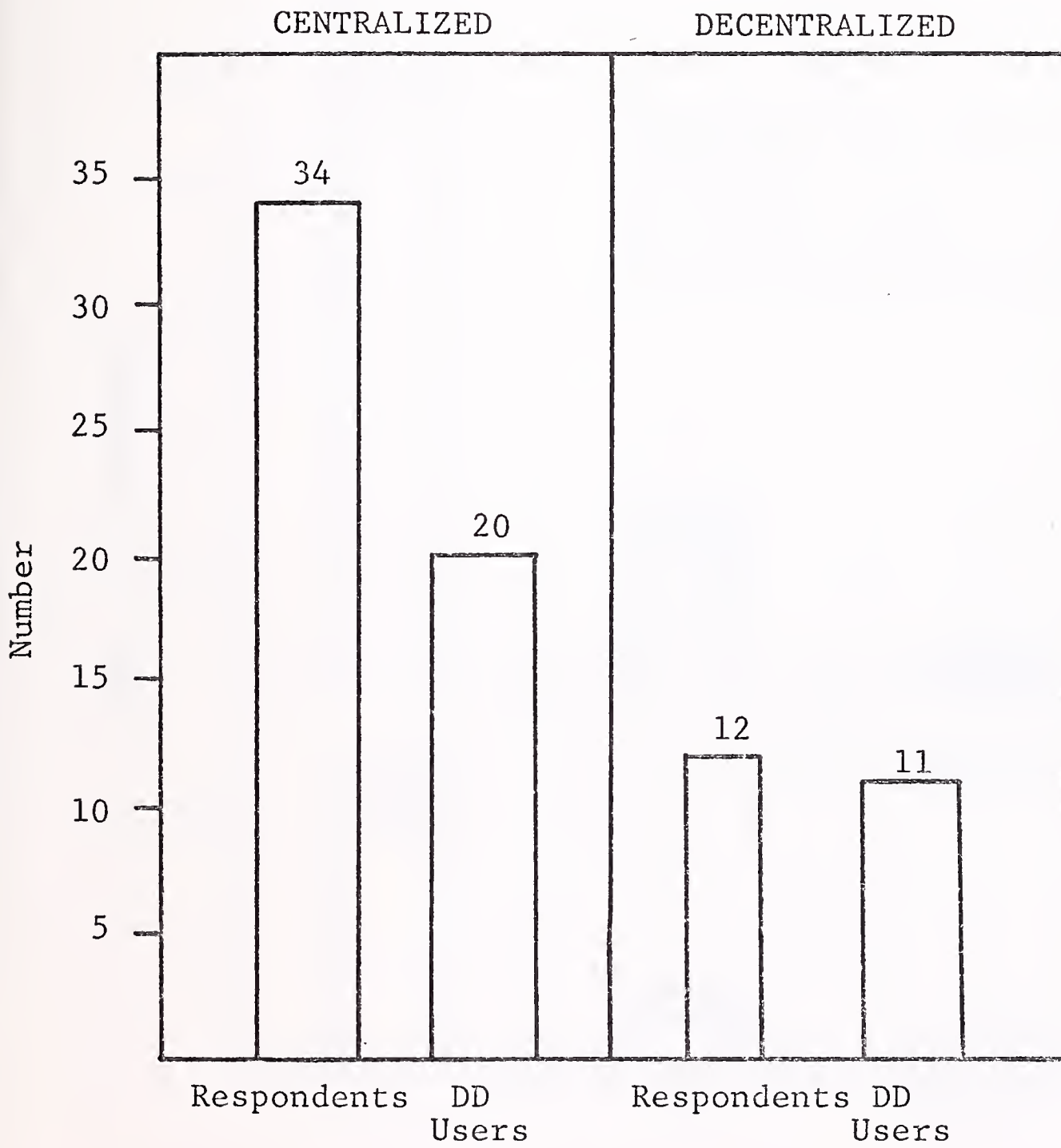


EXHIBIT III-4

DD USE RELATED TO STANDARDS DEFINITION FUNCTION

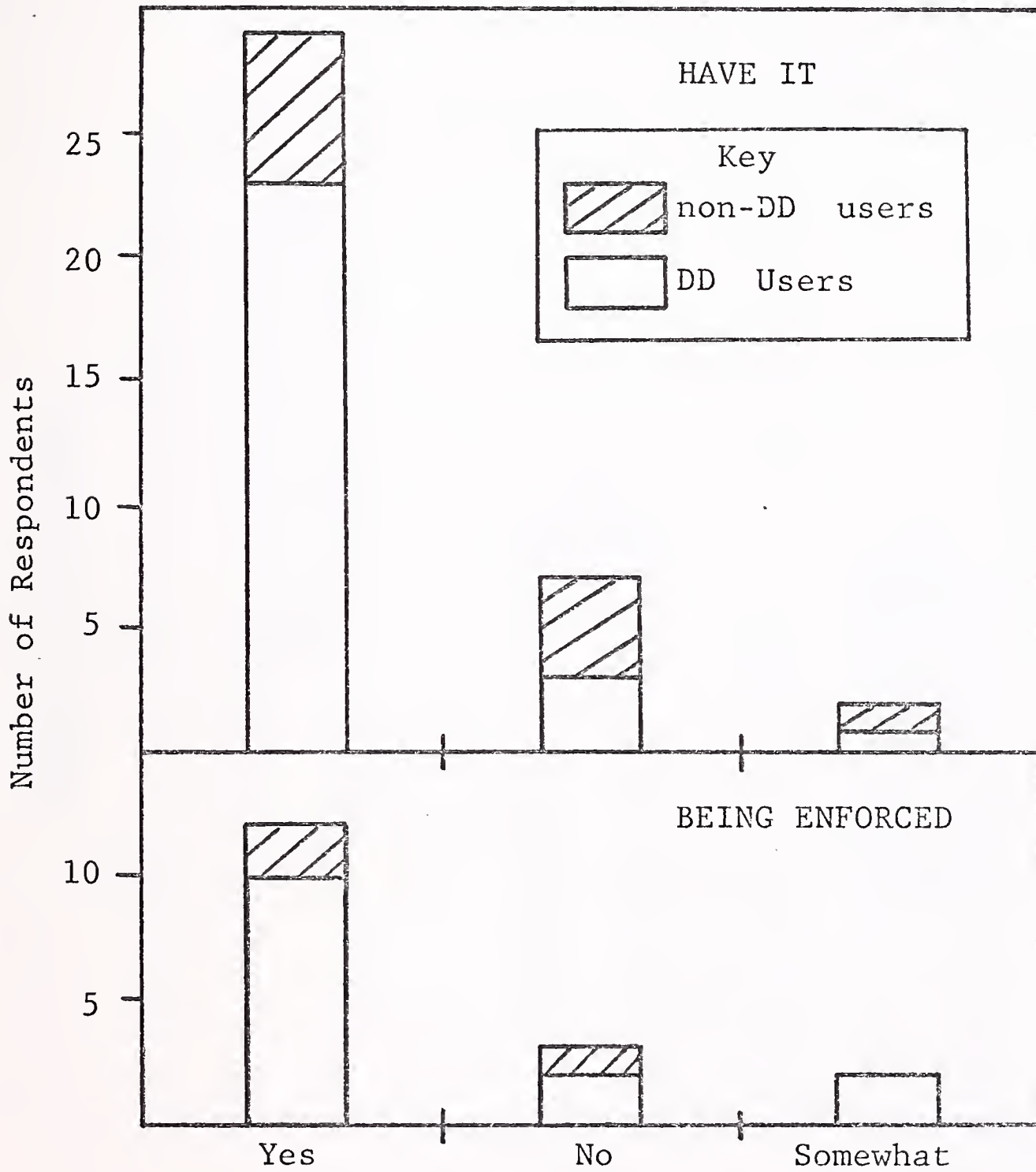
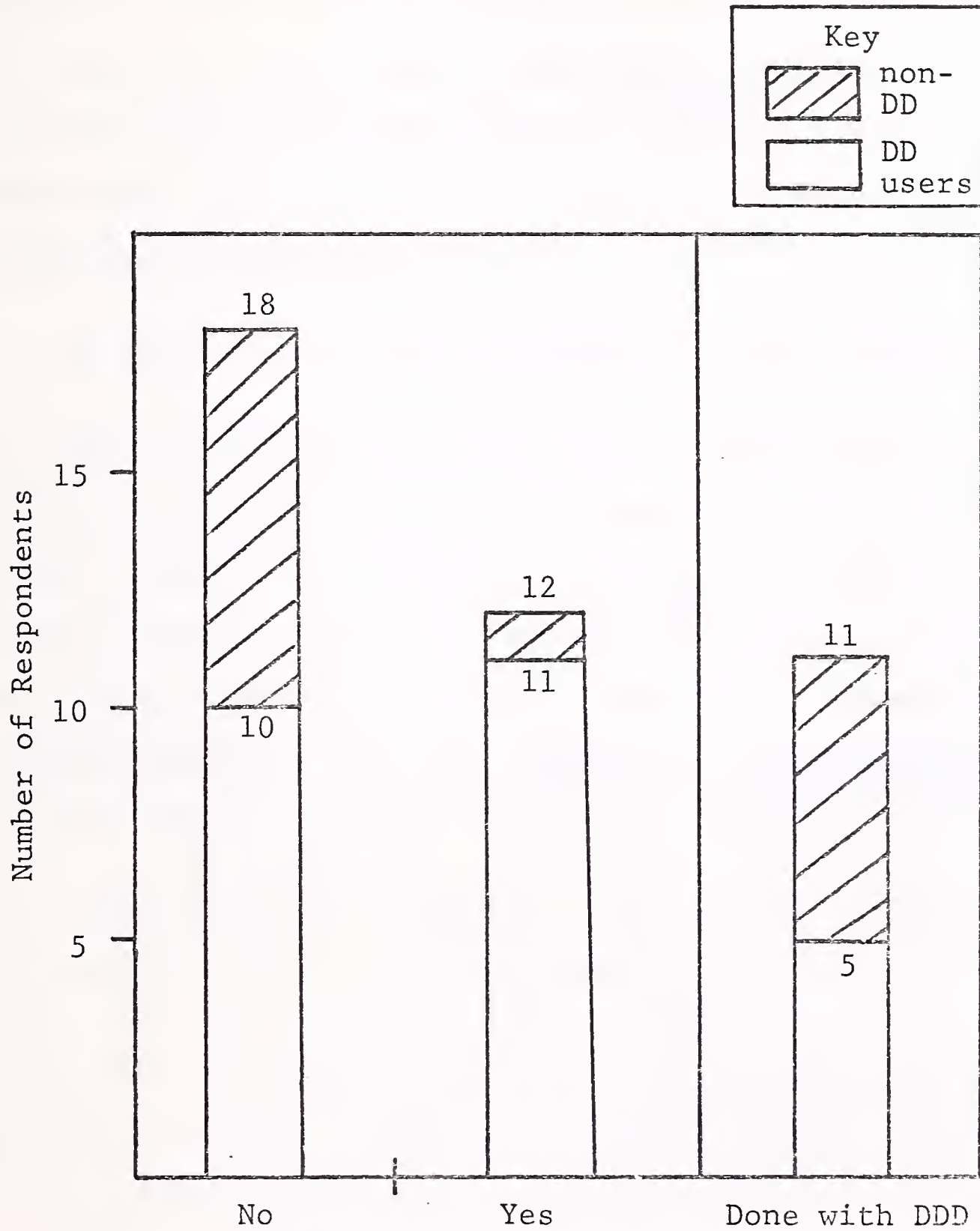


EXHIBIT III-5

RESPONDENTS WITH STANDARDIZED/CENTRALIZED
FILE DEFINITION AND MAINTENANCE



- o It also appears that many of these responding 'yes' to the above question also used the DD as a vehicle in the standardization process.
- o An analysis of the amount of programming time spent on maintenance and development by respondents shows that there is no difference between DD users and non users. In Exhibit III - 6, the average of respondents' allocations of programming is given.

Of the maintenance figure, probably 35% is due to enhancements.

- o Several respondents indicated that data dictionaries had played, or would play some role, in reducing the proportion of effort spent on maintenance. However, since maintenance perforce increases with every new system implemented, obtaining actual performance impacts, if any, of DDs are difficult to measure. Also, since a DD is often accompanied by increased emphasis on standards, establishment of data management functions, and other changes, separating change due to a DD may be well nigh impossible.

In this context, it is significant that not one respondent could come up with any quantifications of the value of a DD.

- o As expected, DBMS users were far more likely to be using a DD than non-users, as shown in Exhibit III - 7. Again, however, the survey is skewed so that it should not be expected that 75% of all DBMS installations have a DD installed.

- o The percentage of IMS users using a DD in the sample was almost identical with that of other DBMS users using a DD. Based on the limited sample, it can be deduced that DBMS users are equally likely to have a DD, regardless of the actual DBMS selected.

EXHIBIT III-6

RESPONDENTS DD USE BY PROGRAMMING
EFFORT IN MAINTENANCE

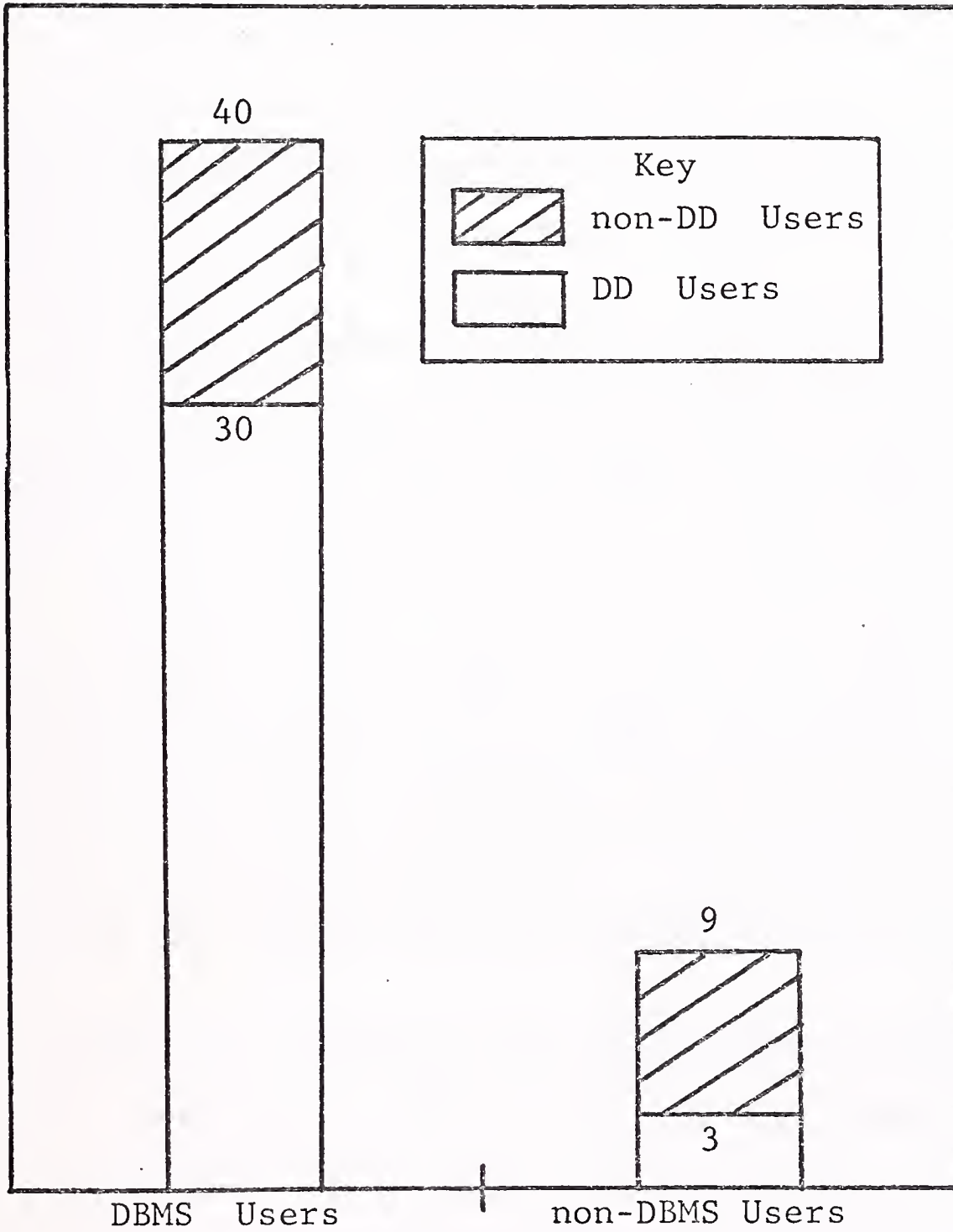
	DD Users	Non-DD Users
Large Respondents	54%	56%
<hr/>		
Very Large Respondents	53%	53%

Percentage given is the percentage of programming effort spent in maintenance. Maintenance is defined as:

- a) Work done to bring a system up to original specifications
- b) Enhancements to existing systems which do not require major rewrites.

EXHIBIT III-7

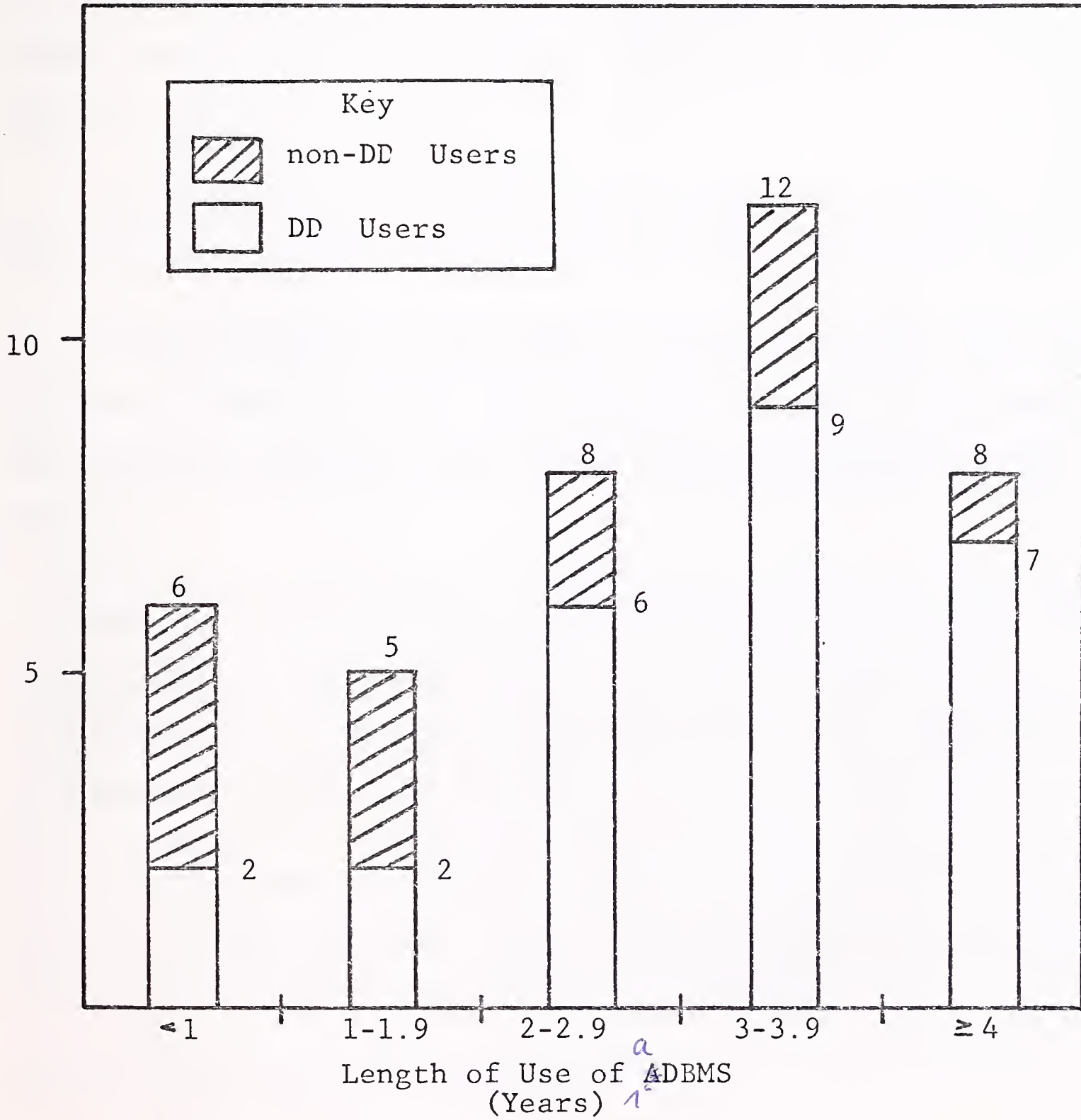
RESPONDENTS' USE OF DBMS AND DD



- o A significant factor in this evaluation is the number of users with their own, often rudimentary, data dictionary. Granted TOTAL users have not had available a commercial package similar to UCCIO and IBM's DD; they seem to have built their own.
- o In the analysis of use of multiple use of DBMS, it appears that they often coexist in installations. Each respondent reporting multiple DBMS in use had them in the same location, albeit on different computers.
- o There were seven respondents reporting use of a DBMS in multiple locations. Of these, five users of IMS also had DDs; the other two, users of ADABAS and TOTAL did not. However, the IMS users also had had a DBMS for at least two years and the others were recent DBMS acquirors.
- o As shown in Exhibit III - 8, there is a very clear relationship between use of a DD and length of use of a DBMS. In fact, the one user of a DBMS for more than four years which did not have a DD, stated that they had realized the need from the beginning but there was nothing suitable available at the time and they had not got around to looking again.
- o Of the DBMS users interviewed, half of them stated they were committed 'very heavily', 'heavily', or '100%' to the DBMS for new development. Of the eight respondents with a DBMS under test, four were similarly committed to its use, and two of the others were committed through price, use of another DBMS.
- o Two installations that had had TOTAL (for three years and four years respectively) were switching to IMS. Both of them had DDs (one has UCCIO and the other Data Catalog). Another TOTAL user with a 155 stated they

EXHIBIT III- 8

USE OF DD BY LENGTH OF USE OF DBMS



would probably get IMS when they moved to a 158. There was no evidence of any movement away from IMS.

- o Based on responses from the survey, it appears IMS will be the standard DBMS for very large IBM users. TOTAL will be used at the large IBM user level.
- o There were almost 60% of DBMS users who responded were using an on-line function. Everyone of these had a DD or was planning to get one.
- o Of the 12 respondents using a DBMS and planning to change, five were upgrading to IMS/VS from IMS, another was going to IMS/DC and the remainder had insignificant changes, other than the TOTAL user planning to change to IMS.
- o Only 9 of the 49 respondents to the survey did not have a DBMS. Of these, all but one plan to get a DBMS: the remaining 'hold-out' has a file system based on IBM's MRP for its major applications. Another is only considering bringing RAMIS in-house from a timesharing-vendor.

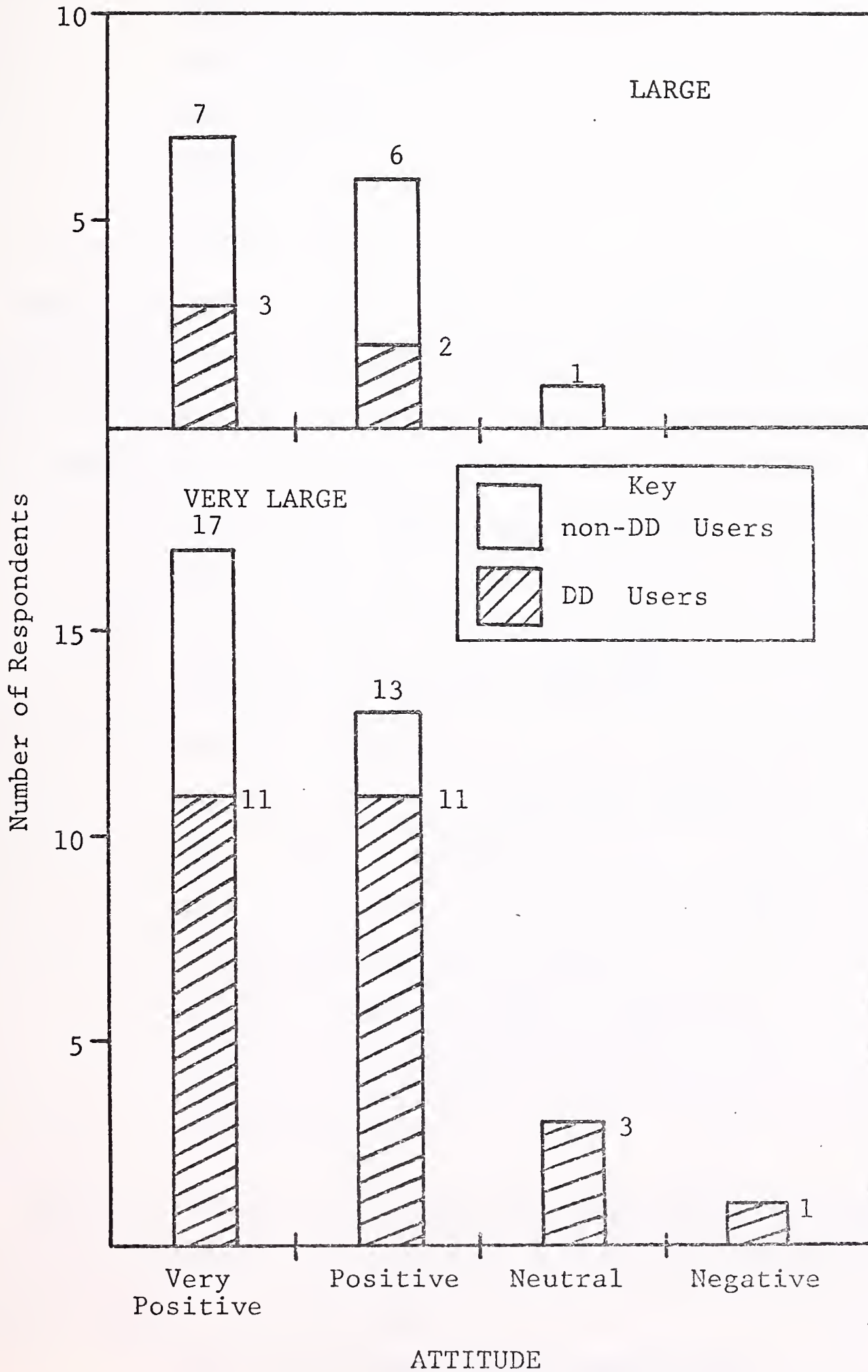
Of the remainder, one intends to build its own system first and then look later at DBMSs available, one has selected IMS/DC, and the remainder will evaluate all DBMS available with required features.

Analysis of The Use of Data Dictionaries

- o Half of the respondents felt very positive indeed about the need for a DD, as shown in Exhibit III - 9. Responses were littered with

EXHIBIT III-9

ATTITUDE OF RESPONDENTS TO DD



phrases such as:

'Must have a DD'

'Absolutely necessary'

'Must have one'

'Essential'

'They're a must'.

o However, it appears that the experience of some DD users has not lived up to their expectations. The problem in analyzing attitude responses, is that actual users respond to their experience rather than their requirements. Some, however, addressed this problem as follows:

'Theory great: haven't seen a good one yet' - IBM DD user counted as neutral.

'Believe in them: current ones leave something to be desired - UCCIO user counted as neutral

'Necessary - need to be improved' - IBM DD user, counted as positive.

'They are necessary; not a good one available' - IBM DD user counted as positive.

'Positive; would have bought if found one (which met requirements)' - non-user counted as positive.

o There is a gap, therefore, between users requirements and packages currently on the market. One criticism expressed by the only respondent counted as negative was as follows:

'They are cumbersome; require redundant data

definition; do not lend themselves to adequate control; too much manual intervention; with IBM or UCC would require 2 full time people.' - respondent using own DD.

Several other respondents felt a DD was a lot of work, but necessary.

o Exhibit III - 10 shows the use and intended use of DDs by the large respondents.

One of the existing DD users with Lexicon will consider changing; while another has a preliminary DD of its own and will get another one.

One respondent (an IDMS user) which has considered a DD will not get one now, but will consider again in future. It stated that none of the existing packages met its requirements; Data Catalog came closest.

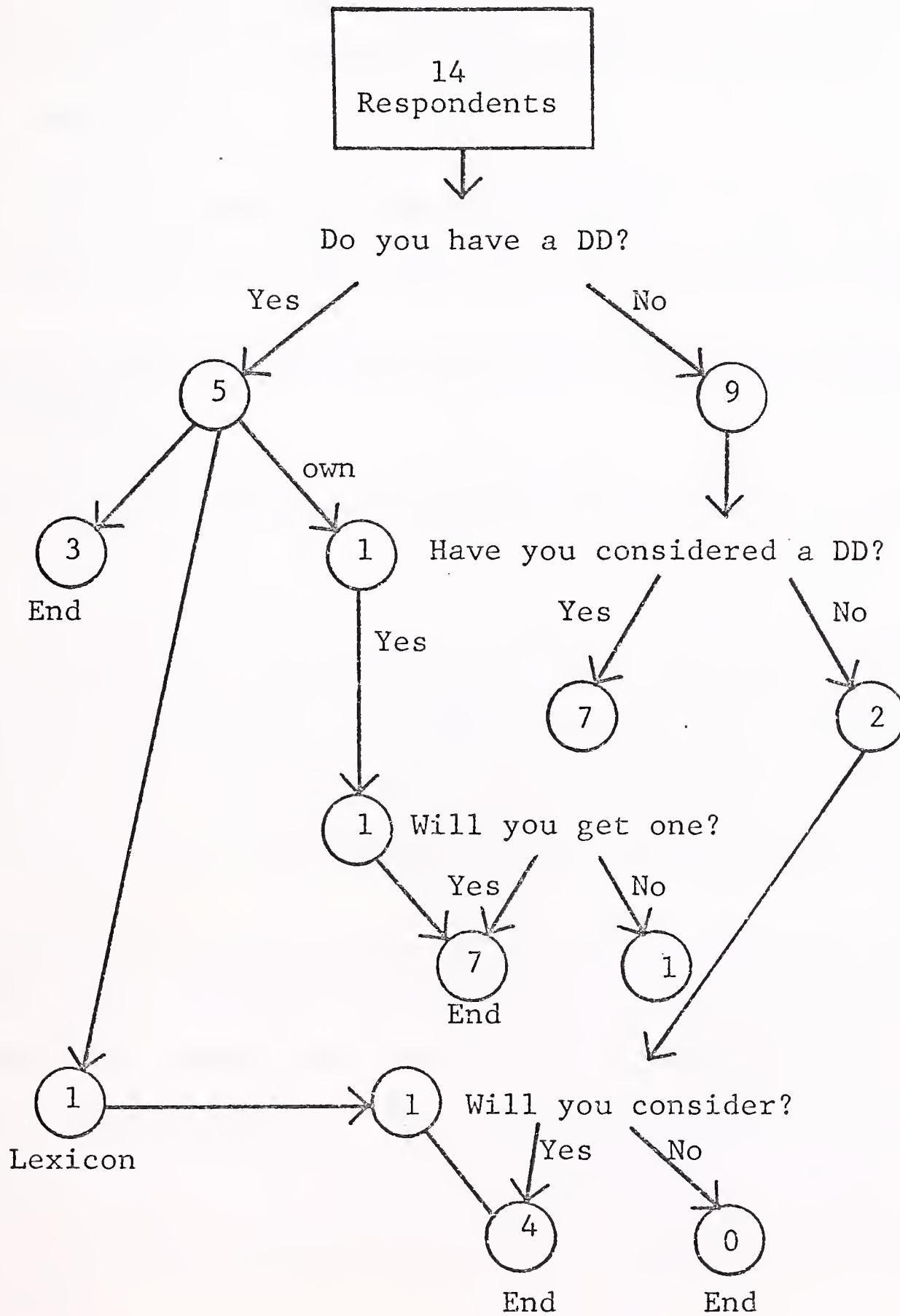
Another respondent using TOTAL and developing its own DD had considered the available products but decided that none of them was good enough. It had examined Data Catalog, Lexicon, UCCIG (eliminated because it was IMS based), and Eastern Airlines.

A respondent with its own DD considered that developing a DD using a DBMS was an ideal way of getting to know the DBMS. This was repeated by several very large users developing their own DDs.

One of the two existing DD users not discussed above was a user of Lexicon as a result of a major Arthur Anderson development project.

EXHIBIT III- 10

LARGE RESPONDENTS' DD USE AND INTENTIONS



- o Exhibit III - 11 shows the intention of those large respondents which stated they will get a DD.

Of the 14 respondents in this category, two were Lexicon users and three others had had presentations from Arther Anderson. Obviously a strong marketing campaign by AA.

- o Of the four large respondents which will consider a DD in future, two of them will decide on a DD at the same time as they select a DBMS.

- o Everyone of the large respondents will end up with a DD according to these responses.

- o Exhibit III - 12 shows the use and intended use of DDs by very large respondents.

One existing user of IBM's DD does not have it operational and will consider another one in future. It requires too much memory. They previously had UCCIO (an early version) but sent it back because they could not make it work. They have participated in a specification project.

Two users of their own in-house DD have considered acquiring a DD from outside. One of them will get one but the other will not get one as they are now. The problem is not price but availability of function: the respondent is familiar with UCCIO, IBM, and Data Catalog and is an IMS user. However, it will consider acquiring one again when it finds one that meets its requirements. In the meantime, it will continue with its own 'or something'.

- o As shown in Exhibit III - 13, only one of the nine respondents stating that they will get a DD has no firm committment or immediate desire to get

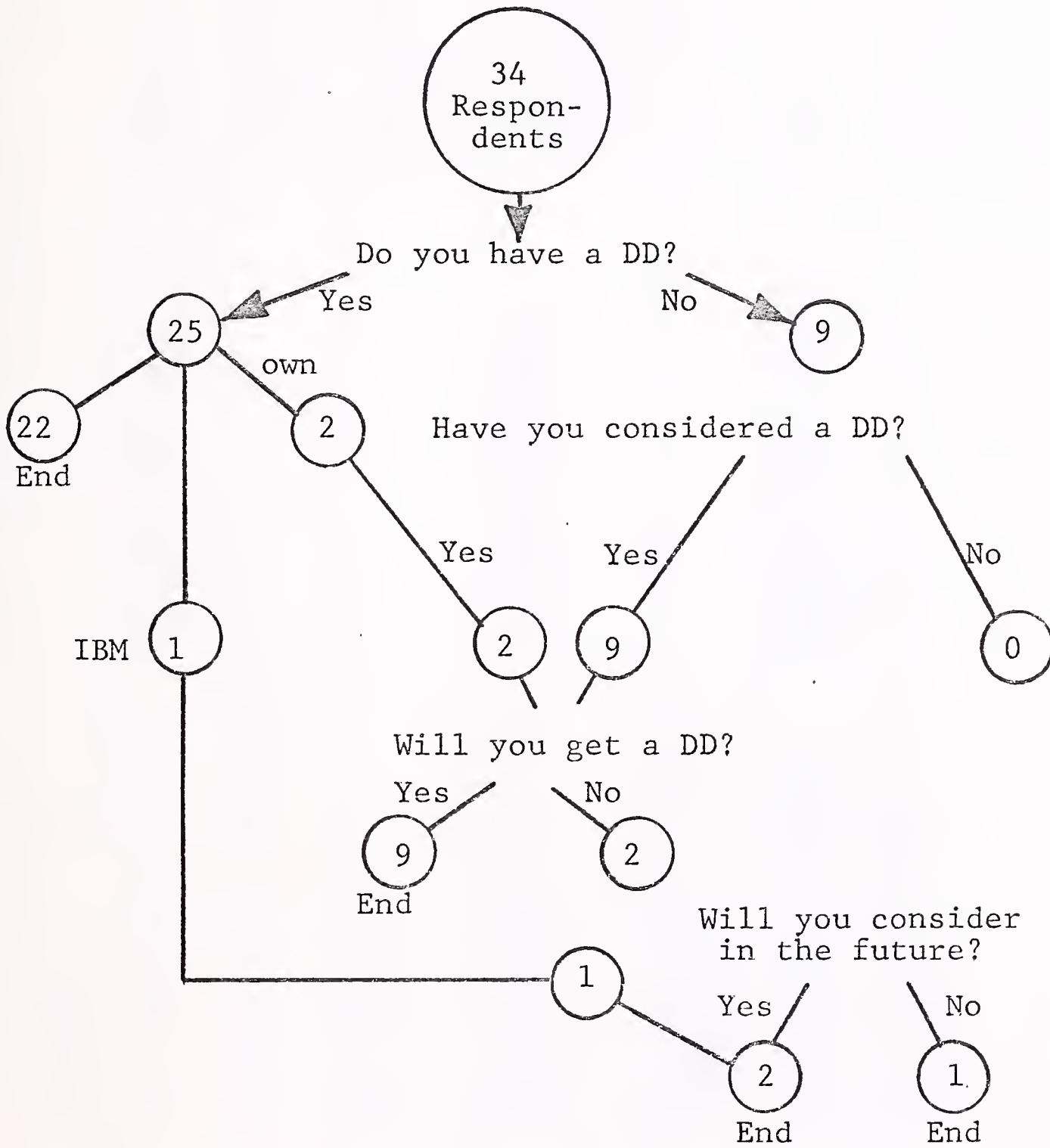
ANALYSIS OF LARGE USERS WHO WILL GET A DD

# of Interview	Commitment	Familiar with DDs	Comment
07	Project Team to evaluate - will get one in 1976	UCC 10, Datamanager, Data Catalog, IBM, Lexicon (AA made presentation)	Don't want to pay consulting fee to AA to set up DBA function with Lexicon
52	Probably get one in '76	Early stages of evaluation	Get one because of pending legislation for public reporting requirements
25	Developing own	Data Catalog, Lexicon UCC 10, Eastern Air.	'None good enough' (Total DBMS user)
2	No commitment	Looking around	If found one suitable, would get it.
29	'Waiting for Cincom'	UCC 10, Eastern Air. Cincom	UCC 10 great; but it's IMS and they are TOTAL
5	None	Had Lexicon presentation from AA	Will get one within six months
59	Data Catalog selected	Pride-Logic, UCC 10, Lexicon, Data Catalog, IBM	Waiting for OS version and their OS system-No DBMS

AA = Arthur Anderson

EXHIBIT III-12

VERY LARGE RESPONDENTS' DD USE AND INTENTIONS



ANALYSIS OF VERY LARGE USERS WHO WILL GET A DD

# of Interview	Commitment	Familiar with DDs	Comment
30	Yes, will develop in-house		
44	IBM delivery in 11/75	UCC 10, IBM, Data Catalog, Data Manager	'Because it is needed!'
45	In process of selection		Delivery around January, 1976
46	No commitment		Will look at seriously before end of 1975
49	80% committed for next fiscal year	IBM, Data Catalog, All ICP advertisers	Leaning to IBM because of support & experience, budget problems now
28	IBM proposed, 1 person assigned		Need ASAP, IMS user
54	Have budget request, no firm commitment	UCC 10, Data Catalog IBM	IMS user. 'Already Late'
50	In principle-none selected		Agency DBA comm. selecting now for delivery in 1976. Use of ADABAS eliminates most needs for a DD
32	No Commitment		Will get with DBMS when select it

a DD. Of the other eight, three are committed to or leaning strongly towards IBM, and the others have not yet made their selection.

o Only one of the 34 respondents will not end-up with a DD according to the responses received.

o Respondents using a DD tied to a DBMS were also asked if they would consider a generalized DD in addition to the specialized one, or if they would consider expansion of their existing DD to cover all data. As shown in Exhibit III - 14 , most users responding preferred expansion of their specialized DBMS to all data. However, three respondents would consider a generalized DD in addition to their specialized one, and three would consider replacing their current DD by a generalized one (of these, one is using Data Catalog, one using its own, and a third using IBM).

Market For A Generalized DD

o Exhibit III - 15 shows the probability of respondents buying a DD such as DATAMANAGER in a competitive situation.

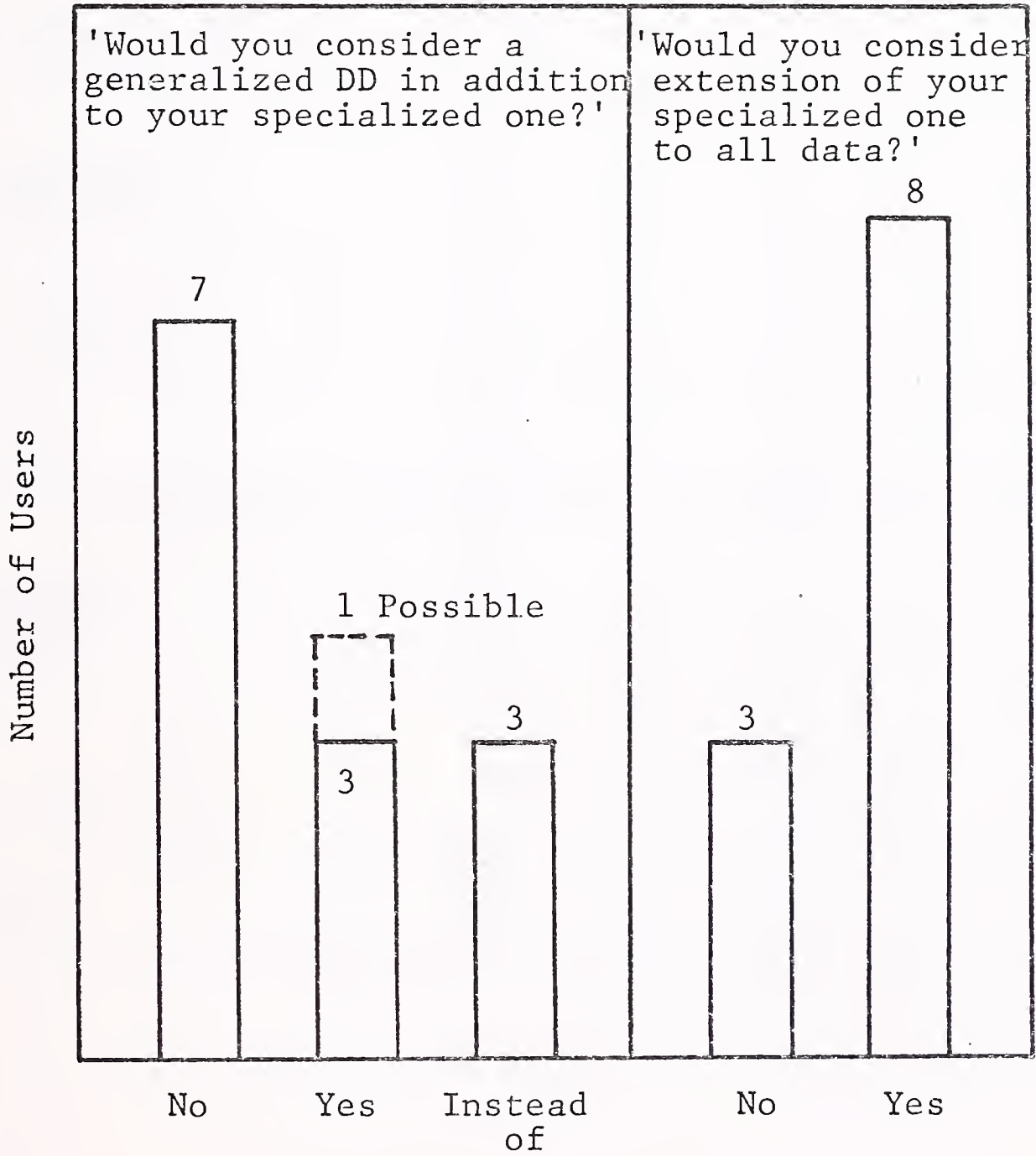
o The respondents included in this table are those which have stated they will:

- get a DD,
- consider getting a DD,
- consider a generalized DD as well as, or instead of, a specialized DD.

o Criteria for inclusion in the 'very high', 'high', and 'lower' portion of the table are:

EXHIBIT III-14

RELATIONSHIP OF CENTRALIZED AND SPECIALIZED DDs



(For respondents whose DD is only used with a DBMS)

EXHIBIT III-15

PROBABILITY OF RESPONDENTS
BUYING A CENTRALIZED DD

	Very High			High			Lower		
				Year					
Respondents	76	77	78	76	77	78	76	77	78
<u>For Data Manager</u>									
Large	2			1			4	1	1
Very Large				1			10	1	1
Total	2			2			14	2	2
<u>For Generalized DD w/DBMS Interface</u>									
Large	4			3	1	1			
Very Large	2			4	1		5		1
Total	6			7	2	1	5		1

- expressed preference for a specialized or generalized DD
- degree of committment
- whether the respondent is developing a DD in-house

o Also shown in Exhibit III - 15 are the probabilities of respondents buying a DD which is generalized, i.e. is not tied to any DBMS, but which has interface with DBMSs. These interfaces, which may be options, provide for internal interaction between DBMSs and the DD; in other words these are automated interfaces which do not require manual or punched card input/output.

o Exhibit III - 16 shows the projected growth of the main DBMSs and DDs within the respondent size through 1979. This is based on a supposition of no new, dramatic conflict with existing products. Assumptions are:

- Both IBM and UCC enhance their products
- IBM introduces a stripped down DD for large users.

o Based on the probabilities given in Exhibit III - 15, the penetration of DATAMANAGER and the hypothetical generalized DD will be as shown in Exhibit III - 17.

EXHIBIT III- 16

FORECAST OF IBM SITES' CHANGES
1976-1979

	Jan. 1 1976	Jan. 1 1977	Jan. 1 1978	Jan. 1 1979
<u>Very Large</u> # of Sites*	750	850	1000	1100
<u>DBMS</u>				
IMS	450	600	850	1000
TOTAL	300	350	355	400
<u>DD</u>				
IBM**	120	200	300	500
UCC 10	120	180	250	300
DC	30	80	130	200
<u>Large</u> # of Sites	2750	3000	3300	3500
<u>DBMS</u>				
IMS	150	300	600	1000
TOTAL	400	600	900	1200
<u>DD</u>				
DC	20	120	150	250
IBM**		50	100	150
UCC 10	10	30	50	100

* Often with multiple installations per site.

** Integrated with IMS - IN 1977 IBM'S DD WILL BE INTEGRATED WITH DBMS

EXHIBIT III- 17

POTENTIAL PENETRATION BY GENERALIZED
DATA DICTIONARY

	1976	1977	1978
<u>Large Respondents</u>			
Incremental DDs	170	200	300
Datamanager	38	44	67
Hypothetical DD	97	114	210
<u>Very Large Respondents</u>			
Incremental DDs	190	220	320
Datamanager	10	11	16
Hypothetical DD	38	44	64
Total Incremental	360	420	620
Total Datamanager	48	55	83
Total Hypothetical DD	135	158	274

Numbers of Installations

At \$15,000 per DD:

Total Market	\$5.4	\$6.3	\$9.3
Total Datamanager	0.7	0.8	1.2
Total Hypothetical	2.0	2.4	4.1

IV TECHNICAL REQUIREMENTS

Analysis of Data Dictionary User Responses

- o As expected, the most required features relate to the basic characteristics of a DD, as shown in Exhibit IV-1. Users want data element definitions and the ability to define relationships between all parts of a program system - including inputs, outputs, files (at various levels) programs, and data elements.
- o The reporting characteristics receive the second highest number of mentions and again the importance of the DD covering relationships is established by the responses. The ability to easily access and retrieve data using scans, sorts, report generators and other tools is frequently mentioned.
- o Although ease-of-use was only singled out five times by respondents, this attribute was present in many responses categorized elsewhere. It appears that users do not find present systems easy to use: in particular reports and displays from the present DDs are not easily understood by non-DP people. Further, in many instances responsibility for definitions rests outside the DP department - the inputs to the DD should also then be able to be handled by non-DP people. Again this is not now the case.
- o Integration with the actual programs being developed is required, particularly when a DBMS is being used. Control block generation was frequently mentioned as a mandatory requirement. Input from source statement or copy features was also required.
- o Four respondents (two of which were Lexicon users) also wanted validation procedures included in a DD, while another four wanted to be able to handle aliases and synonyms. One of the latter merely wanted to trap them, for subsequent elimination.

EXHIBIT IV-1

REQUIRED FEATURES OF A DD

# of mentions	Basic Features	Some Comments
19	Data element definition and relationships	<ul style="list-style-type: none"> o Text description required o Item level documentation o Define hierarchy of relationships
16	Retrieval and reporting characteristics	<ul style="list-style-type: none"> o Cross-reference lists (6 mentions) o Where used reports (3 mentions) o KWIC List and scan capability o Report Generator(intelligent output)
7	Generation of control blocks or similar program code	<ul style="list-style-type: none"> o 'Create DBDs and DSDs through the DD-IBM does not do this o Generate storage for file control
7	Integration with a DBMS	<ul style="list-style-type: none"> o Interface with IMS-but be independent of it for operation o Interface with any special DBMS
5	Ease of use	<ul style="list-style-type: none"> o 'Enter data only once then use it' (IBM DD user) o Readability by non-DP people (several respondents wanted reports to be useable by end-users)
4	Validation procedures	<ul style="list-style-type: none"> o Validation in machine processible form
4	Aliases and Synonyms	<ul style="list-style-type: none"> o Trap synonyms
<u>Other</u>	<ul style="list-style-type: none"> o On-line inquiry(update) (3 mentions) o Operating System Support (2 OS and 1 DOS and OS) o Security (2 mentions) o Well Documented o Test Data Generator o Specification responsibility included in DE definition 	<ul style="list-style-type: none"> o Non-computer data (3 mentions) o Internal and external formats o Not too much abbreviation o Standard mnemonies o Powerful-lot of processing for little /input o Generalized o Documentation system for analysis and programmers.

- o Of the other required features, the on-line inquiry and update feature reflects the problems that many users have with the magnitude of the effort required in running a DD. Also, three users specifically singled out the ability to handle non-computer data and several others alluded to it.
- o The responses as to features required were probably heavily dependent on the characteristics of the packages being used. Thus 'ease-of-use' may become a 'required' feature for a user of IBM's DD or UCC 10 which appear to be difficult to use, whereas that is simply a 'good' feature for DATA CATALOG.
- o Exhibits IV-2 to IV-5 show the responses of users to questions on the features they found 'missing', 'poor', or 'good' in their DD.
- o As can be expected, respondents with their own dictionaries found less features missing or poor than those that were good in their DDs. In the package cases, however, the balance lay the other way, culminating in UCC 10 which had 19 poor or missing features quoted by 5 respondents and only 1 good feature identified.
- o Emphasis on the ability to get data in and out of the DD runs through all the DD comments.
- o The most striking area of concern was reporting in the case of IBM DD users. All five of the respondents with anything to report in this area gave the IBM DD poor marks for its reporting characteristics.
- o Exhibit IV-6 shows the respondents' reaction to features suggested to them. Responses including 'mandatory', 'nice', 'good', etc., have been classified as '+ve'; responses which are classified neutral or 'ok' are 'yes', 'ok', 'might be +ve', etc.; '-ve' responses are 'no', 'not necessary', etc.
- o In addition to these features tabulated, the following features were suggested:

UCC 10 FEATURE ANALYSIS

Features Missing or Poor	Features Good
<ul style="list-style-type: none"> o 'Nothing Known' or 'Not Important' o Data Base Design Aids-Better Support for Design Activity o Table Management System especially for corporate tables outside DP-badly missing o Project Management/Control System (2 mentions) o Change Management/Control capability (2 mentions) o Relationships of data elements/files/programs o Other data description items commented on: o Characteristics should be at intersection level rather than element level o Need location of data by site/host/file o Handle non-IBM data elements for distributed processing o Other reporting comments: o Hard to keep track of reports-need to go back to 'fields' for duplicates o Keyword features need improvement-need 'and/or' keywords o All UCC packages should interface with it-would have bought UCC 7 if it had o Documentation could be better 	<p>Output facility to interface with IMS best I've seen</p>

IBM DD FEATURES ANALYSIS

Features Missing or Poor	Features Good
• None that are necessary	• Lot of capability for defining relationships
• Reports are poor:	• Interface with design aids
Ability to get everything out is not complete	• Design and flexibility-potential is there
Readable reports missing	• Data structure (good)
Incomplete reports	• Aliases very good
Reporting is poor	• Generalized IMS control-block features are automatic
Output reports (are poor)	• File design one of the best
Control and reporting (are poor)	
Reporting function is poor	
Reports (are missing)	
• Input is cumbersome-very critical	
• External translation of data element description	
• Abilities to handle COPYLIB input and general tables are missing	

Respondent Comments

EXHIBIT IV-4

FEATURES ANALYSIS FOR USERS WITH THEIR OWN DD

Features Missing or Poor	Features Good
(only 3 out of 9 respondents included)	(5 respondents included)
o Validation definition relating to data on different DBMSs is poor, but..	o Validity criteria are good
o Version control poor	
***	***
o Dictionary does not describe COBOL picture-so, no generation capability	o User describes/defines data elements
o Not used in line with programs so no control	o Key word sorts and indicators
o Not responsive to program changes	o Documentation and reporting
o Updating is a problem	o Tied to DBMS (ADABAS) Data element definitions only entered DD, they are generated for DBMS
***	***
o DD system not designed to support minis - in distributed processing	
***	o Generates data division record structures
	o Checks all compilations for Std. data names

Respondent Comments

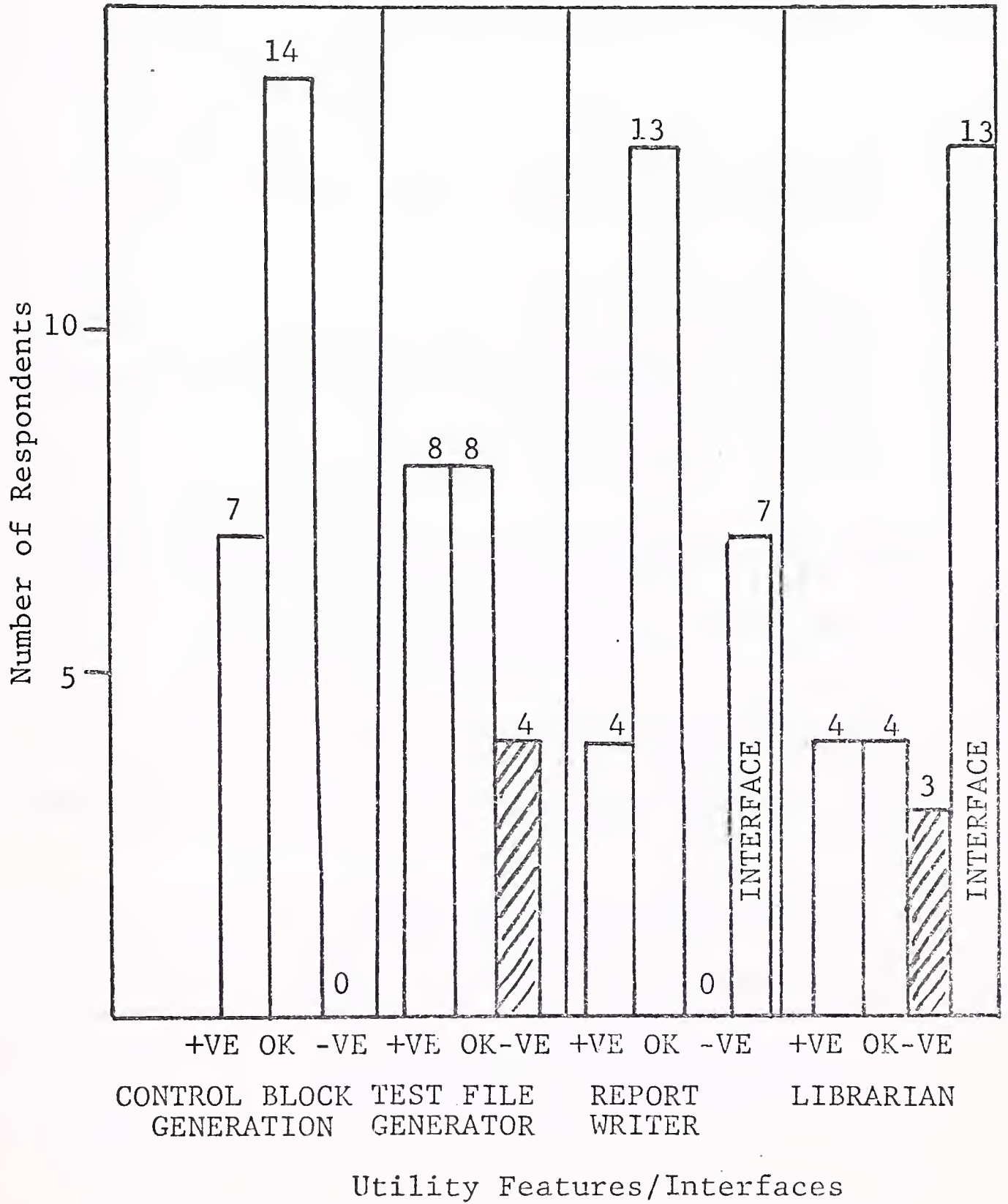
OTHER DD USERS FEATURES ANALYSIS

Features Poor or Missing	Features Good
<u>Data Catalog</u>	
<ul style="list-style-type: none"> o On-Line query capability missing (getting it) o On-Line query capability not available yet o Data elements not well enough documented o Doesn't automatically interface with IMS -punches out cards o ISAM handling (is poor) o Limitations on number of synonyms o Line numbering technique adds to complexity 	<ul style="list-style-type: none"> o Documentation is good-easy to find what you are looking for o Reports are good-easy to use o Can put <u>absolutely</u> anything into it o Forms most important-very good
<u>Lexicon</u>	
<ul style="list-style-type: none"> o No DOS version o Cumbersome-pulling cards for updates <p>(Nothing reported missing)</p>	<ul style="list-style-type: none"> o All Report writer function - but overhead is high o COPYLIB function-so programs are controlled
<u>Pride-Logic</u>	
<ul style="list-style-type: none"> o Three levels of defining data-how it is physically stored and used. May be missing in all DDs o On-Line capability poor 	<ul style="list-style-type: none"> o Tied to way they do business-through Pride.

Respondent Comments

EXHIBIT IV-6

RESPONDENTS EVALUATION OF
UTILITY FEATURES/INTERFACES



- validity checking - twice
- interface with IMS utilities

o The report writer feature and control block generation were generally well accepted. In the case of the report writer, an interface with an outside, standard report writer would probably be better accepted than an internal, non-standard report writer.

o In the case of a librarian, an interface with PANVALET was mentioned by 13 respondents, of which three were only 'possibly' interested while the remainder were positive.

o For documentation and reporting features, Exhibit IV-7 shows that data depiction (graphically demonstrating the relationships of files, records and data groups) has a very high ranking with almost 50% of respondents considering it very positively.

o The teleprocessing environment documentation also achieves high acceptance: somewhat lower than cross reference reporting needs which must be regarded as a mandatory part of a DD.

o Statistics collection and the feature reporting the relationship between T/P and data structure were the only ones with more negatives than positives.

o The most important programming support feature, as shown in Exhibit IV-8, is that of enforcement of programmer use. Only one respondent (a user of IBM's DD) felt it was not practical as a function of the DD. Another respondent considered it necessary to be able to operate outside it if desired.

o In examining the responses on batch and on-line access capability shown in Exhibit IV-9, only 2 of 26 respondents felt negatively about it: one considered on-line too fancy (a user with its own DD) and the other considered on-line 'not important' (a Data Catalog user). Five respondents

EXHIBIT IV-7

RESPONDENTS EVALUATION OF DOCUMENTATION
AND REPORTING FEATURES

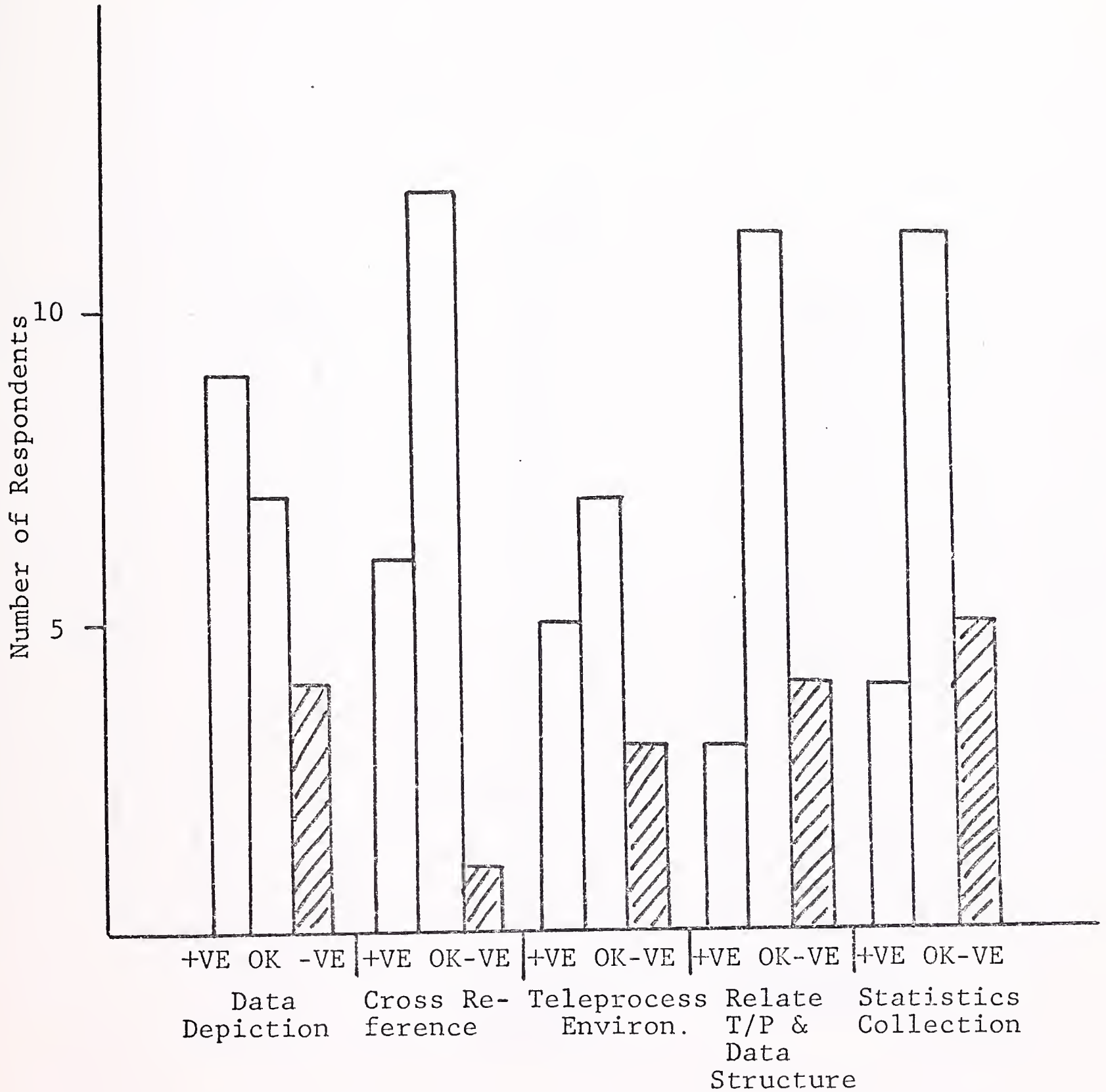


EXHIBIT IV-8

RESPONDENTS EVALUATION OF PROGRAMMING/SUPPORT FEATURES

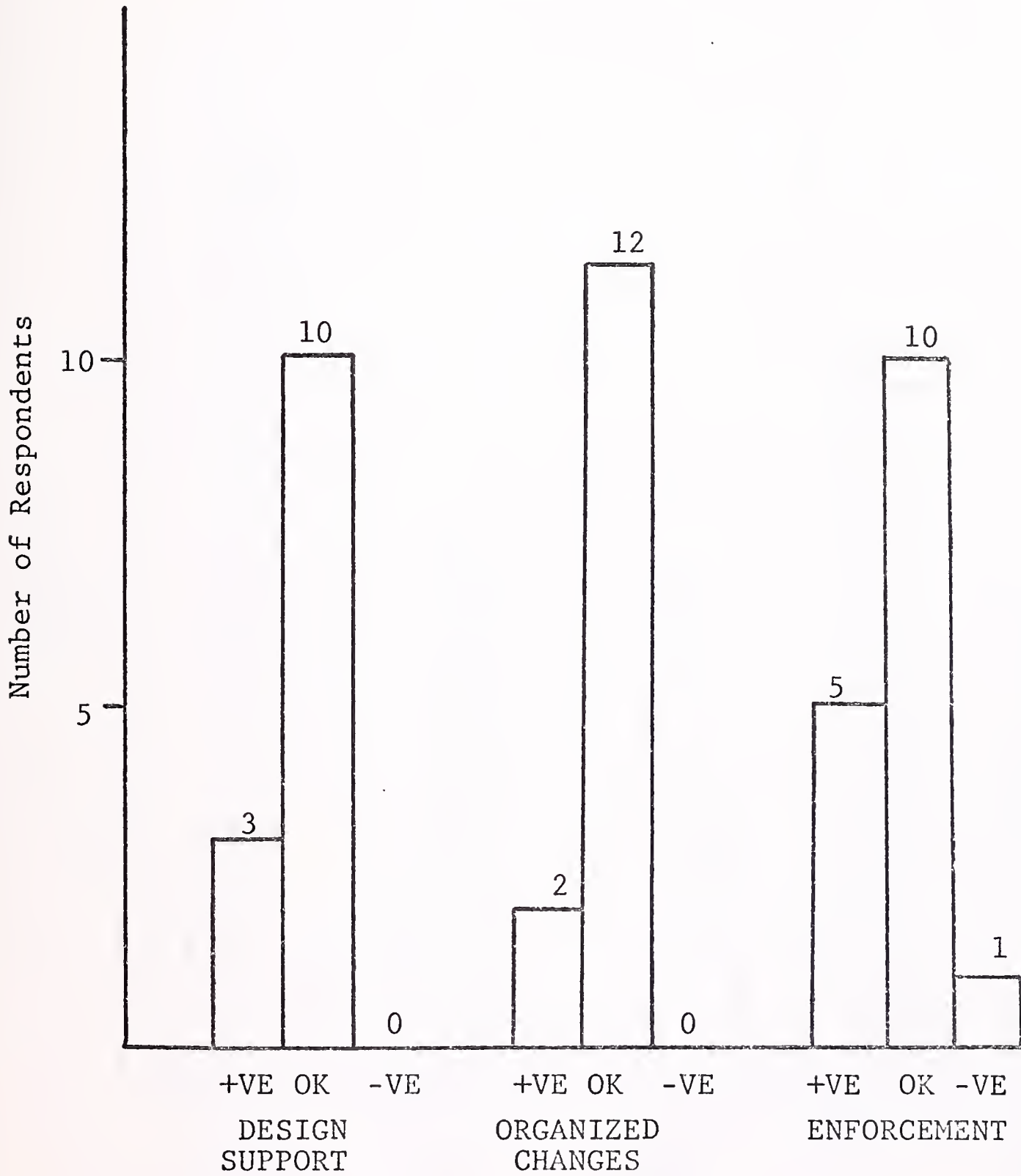
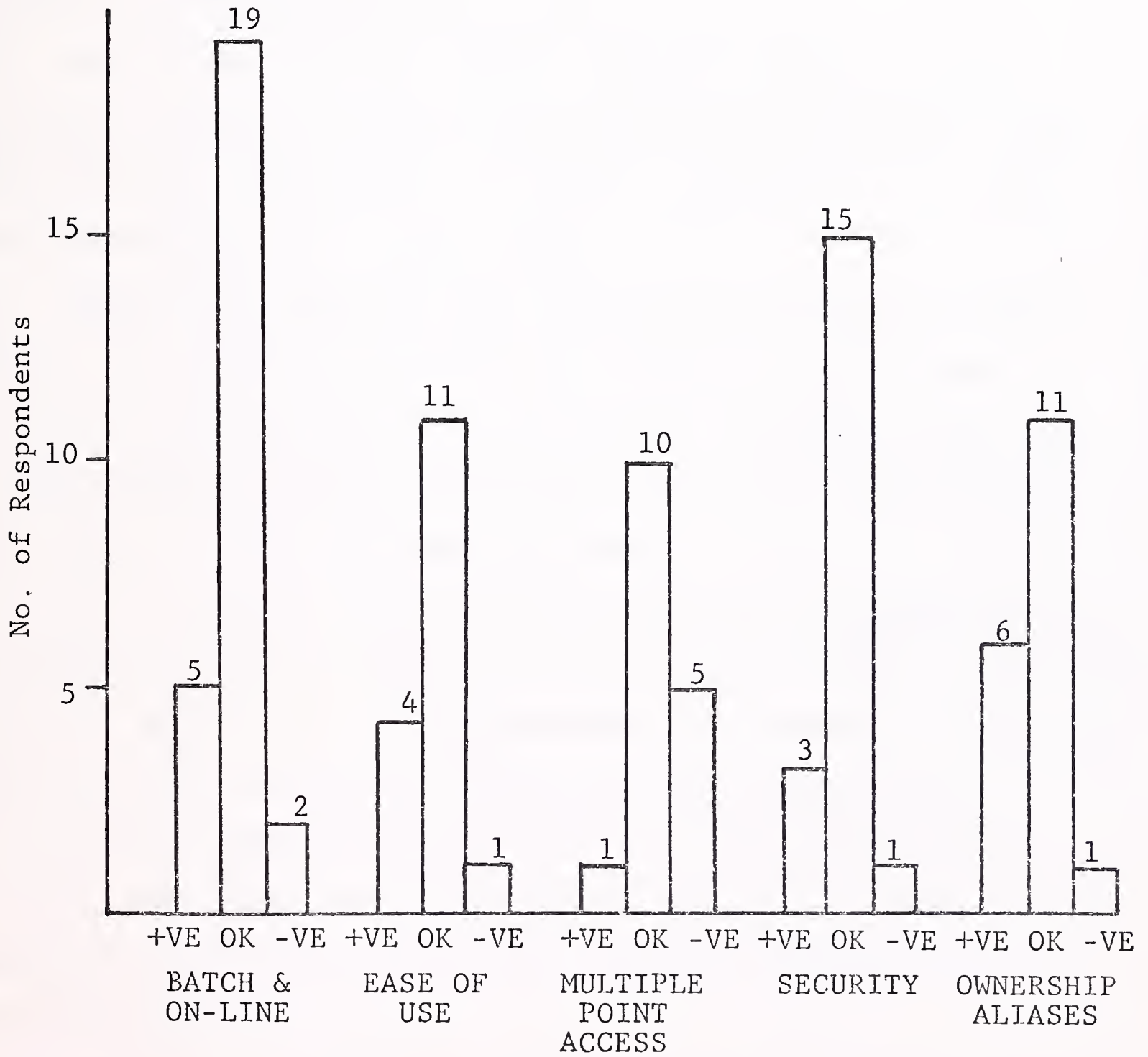


EXHIBIT IV-9

RESPONDENTS EVALUATION OF OPERATIONAL CHARACTERISTICS



specifically considered it positive to have an on-line capability; one stated it would be 'nice to have - since it would save papaer'.

o Multiple-point access was the least important feature, as may be expected since most installations responding were centralized: even of those with multiple sites, it was generally not considered important, probably because each site is independent. One respondent commented that it accessed its own DD through RJE.

o Security was fairly well received, although respondents may have confused security on the DD itself with storing levels of security on the data elements. Two respondents specifically identified the need to capture the security level of data elements.

o Aliases and synonyms were felt to be very important by five respondents: one respondent considered it very important for old systems. This would seem to be generally true, since this requirement is also identified elsewhere. Aliases are relatively unimportant for new development but a must in trying to integrate old systems.

o Of 29 respondents to the question on enforcement of programmer use of a DD which have a DD, 18 would do it by forcing the programmer to go through the DD to get a job done. Three respondents do not enforce it but encourage its use through training and other means as shown in Exhibit IV-10.

o As shown in Exhibit IV-11, interfaces with report writers and on-line software packages were equally required by respondents. The amount of required on-line interface backup the analysis of the on-line feature of a DD being important.

o Respondents using IMS also specifically identified requirements for interfaces with the IMS design aids.

o PANVALET was the package which most often required an interface according to the respondents.

o Of the respondents using a DD, 13 of them specifically stated 'no'

HOW RESPONDENTS ENFORCE PROGRAMMER USE OF A DD

Automatically	Manually	Other
Through DD:		
- Can't write programs otherwise	- Management Policy and Standards	- Training new programmers
- Use COPYLIB function	- Management control	- Administratively for applications program-
- Hook into compilation procedure	- Audits	- mers-systems program-
- Generate IMS environment	- Project Review	- mers have to go through for SysGen
- Via DBDs and PSBs		- Don't enforce packages let them sell themselves
- Integrated with DBMS		- By threat and environment
- DBA's responsibility is COPYLIB		- People use because of sales effort-If enforced it would not be productive
- Programs can't be written outside DD		- Play "You bet your job" Currently not enforced let it sell itself.
- Data division comes from DD		
- Can't get to data base without going through DD		
- Through copy feature		
- Must use COPYLIB		
- Can't get by system - if they do fire them.		
- Don't allow anyone outside of DBA group to update DD		

Respondent Comments

EXHIBIT IV- 11

RESPONDENTS' REQUIRED INTERFACES WITH
SOFTWARE PACKAGES

Mentions	Required Software Package Interface
5	IMS Data Base Design and Performance Aids
4	PANVALET
3	<u>Report Writers</u> ANY
2	MARK IV
2	ASSIST
1	DATA ANALYZER
1	GIS
3	<u>On-Line Software</u> INTERCOM
3	CICS
2	TSO
1	WYLBUR
1	<u>Other</u> COPYLIB
1	TLMS
17	None Required

in response to the question on required interfaces, while 16 required an interface of one kind or another.

o Few serious operational problems were encountered by respondents as shown in Exhibit IV-12. Most of them have been resolved. In the sole case where a user discontinued using its DD, which it had developed, because of problems encountered, it now rates that decision a 'major mistake'.

o Respondents with a DD and a DBMS were asked whether it was an advantage for the DD to be tailored to the DBMS. Of the 33 respondents, 24 felt it was an advantage. Of the nine which considered not to be so, several stated they preferred a generalized package.

o However, it was also indicated that a specialized (tailored) package did not need to run under the DBMS, as with UCC 10 or IBM's DD. In fact since the overwhelming majority of respondents considered that a DD should be implemented before a DBMS it was strongly indicated that its DD should be able to operate independently.

o As shown in Exhibit IV-13, two thirds of the respondents were using their DD for all data, although words such as 'target', 'plan', or 'hope' were prevalent. Of the 14 respondents using it only for a DBMS several also stated that was their initial goal.

o Five of the respondents volunteered that they would use the DD for non-computerized data. This may well be a significant trend.

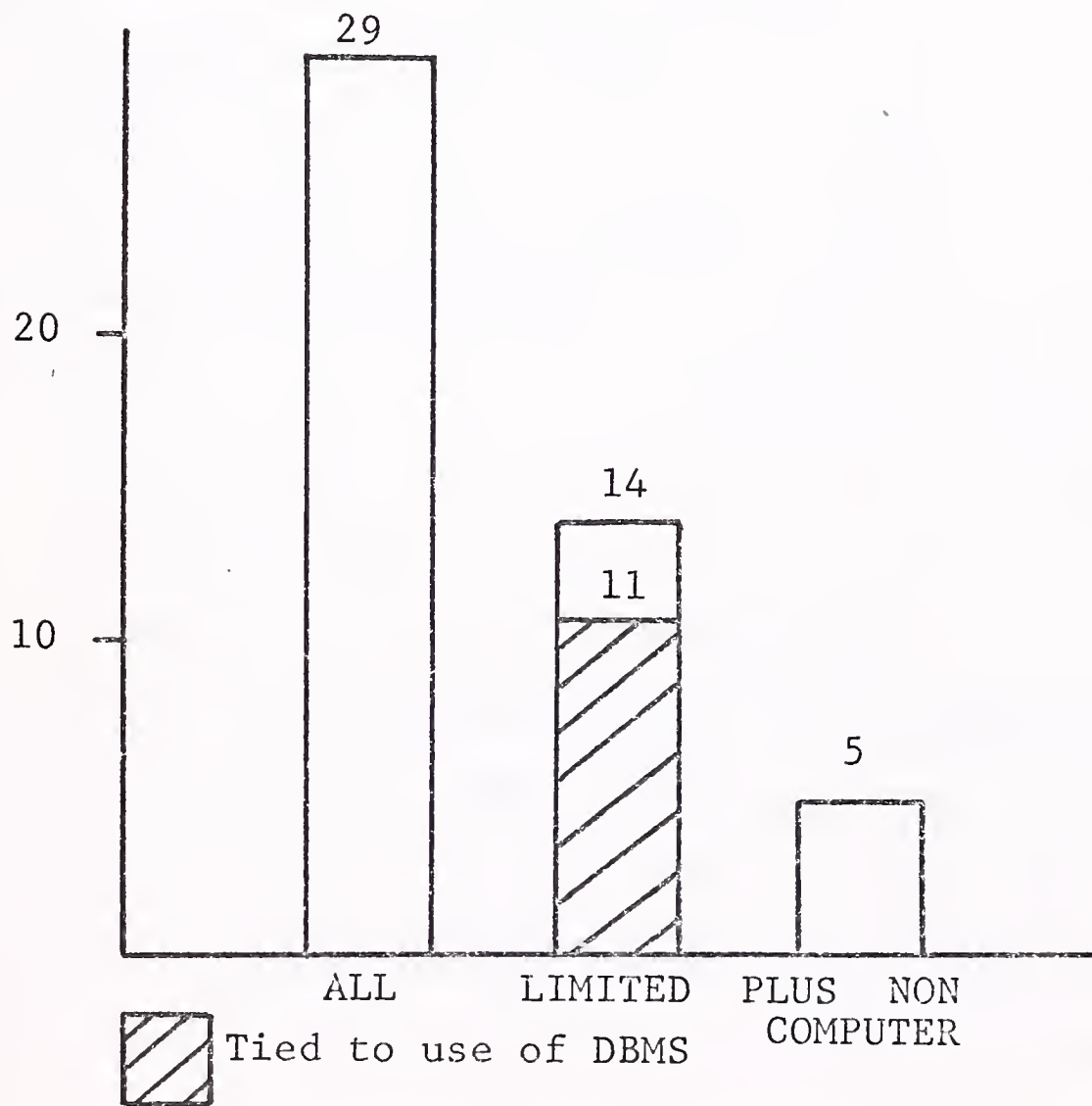
o Few respondents volunteered information as to why they only used a DD with limited data; 'control' of the DBMS was volunteered several times, while two respondents considered their specialized DDs unuseable in the general case. In fact, one respondent had two DDs, one for general data and another for IMS.

OPERATIONAL PROBLEMS ENCOUNTERED BY DD USERS

Operational Problems Encountered	Solution
1. How to separate from IMS problems? IBM DD language is <u>gross</u> : it is stilted. Cannot have the whole system <u>from test to production</u> .	UCC 10 is also <u>gross</u> : worse. Data Catalog is simpler, but not adequate (IBM DD user)
2. Not enough core.	Got more core (early user of UCC 10 VS version)
3. Initially had obsolete ISAM. Some trouble loading Data Catalog.	Got source code (Data Catalog user)
4. It slowed down development, because it was not being used properly.	Stopped using it - a major mistake! (own DD user)
5. Initially problems with Lexicon on IDC because IDC has non-standard operating system	Brought in-house.
6. Problems with package: also getting people to use it.	Package debugged - showed people it would save time (early UCC 10 user)
7. Long running updates due to massive amount of relational data, due to data structure.	Getting used to it (IBM DD user)
8. Input/output poor	Wrote preprocessor and added additional reports (IBM user)
9. Motivating people to use it.	Convince them it's needed
10. Converting systems with good documentation, but wrong format.	---
11. Where to start? Old or new systems?	Chose new systems

EXHIBIT IV-13

USE OF DD FOR ALL OR PARTS OF
RESPONDENTS' DATA



Specification of Required Features of a Generalized Data Dictionary

- o **EASY TO USE.** The DD should be easy to use in the following senses:
 - redundant input of data must be eliminated
 - reports should be readable by non-DP people
 - documentation should be easy to read and reference
 - input forms should be simple and easy to complete, directions for use should be simple.
 - abbreviations should be avoided.
 - o **COMPLETE CROSS REFERENCE CAPABILITY.** The DD should show cross references across and among all system elements including:
 - programs and program elements
 - files, records
 - data elements, at several levels
 - reports and other outputs
 - inputs
 - multiple DBMS
 - multiple locations
 - multiple host computers, including minis
 - o **INTERFACE WITH DBMS AND PROGRAMMING LANGUAGES.** The DD should not be required to operate under any one DBMS. However, it should have optional features which will handle interfaces with each standard DBMS, initially TOTAL and IMS. It should be able to handle multiple DBMS. The user should optionally be able to enforce programmer use through relationships with compilers.
- On the other hand input should be acceptable from source statement code and copy facilities should be present.
- o **VALIDATION FEATURE.** An advantage, but not a requirement, would be to have a data validation capability.

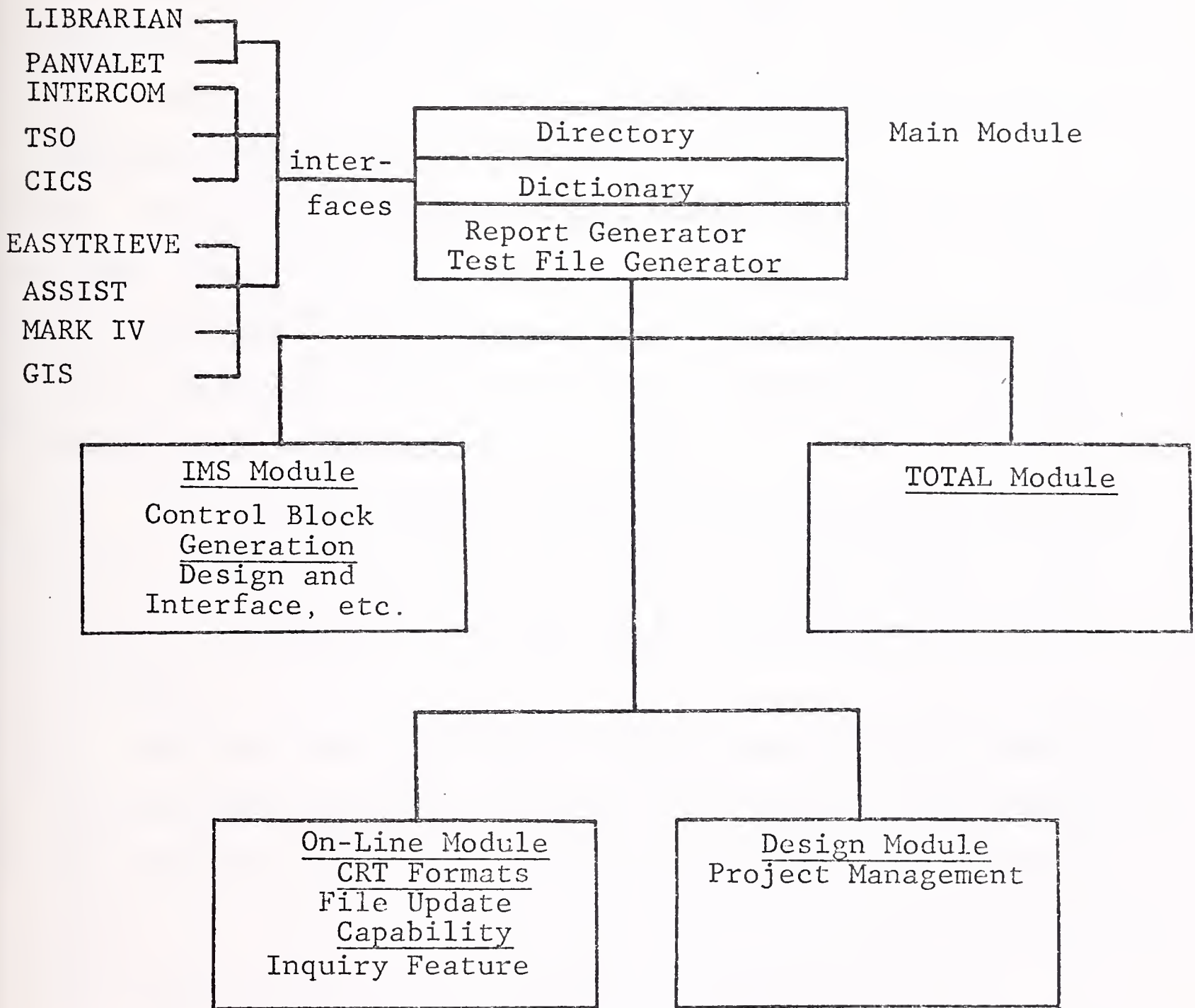
- o ON-LINE CAPABILITY. An on-line inquiry and update capability is mandatory and it should be CRT oriented. However, it should be an optional feature and priced separately.
- o HANDLING OF NON-COMPUTER INTERFACES. The DD should be capable of handling non-computer data. This will require the ability to define and handle non-computer files and tables. Also definition of data elements in both an external (non-computer) and internal (computer) format is required.
- o SYNONYMS, MNEMONICS, AND ALIASES. The DD will handle these.
- o REPORT GENERATOR. The DD will offer interfaces to standard report generator function, such as EASYTRIEVE and ASSIST. However, it should also have its own so that a user will not be forced to purchase another package in order to properly use the DD. The internal report generator can be fairly simple: however, it should include key word in context capabilities and be able to handle Boolean logic key searches ('AND/OR').
- o OPERATING SYSTEM. The DD should be able to operate under DCS, DOS/VS, OS or OS/VS. It should also support multiple operating systems on host computers.
- o INTERFACES. The DD must offer an interface with PANVALET.
It should also interface with the standard on-line software, such as CICS, Intercom, and TSO. These can provide for multi-point access with certain additional security features.
- o PROJECT MANAGEMENT AND SYSTEM DESIGN AIDS. The DD should include, as an option, a project management system and design aids which will enable the DD to be used during all stages of a system design, development and installation phase. The project management system should include a networking feature, such as PERT or CPM.
- o DATA DEPICTION. A capability to graphically display relationships should be present in the DD.

o OWNERSHIP. The DD should clearly be able to handle ownership and responsibility at several levels.

o STRUCTURE. The DD should consist of a set of modules which as user may purchase separately, as shown in Exhibit IV-14. UCC should be able to migrate UCC 10 users to it.

EXHIBIT IV-14

UCC 10 - GENERALIZED DATA DICTIONARY



V MARKETING REQUIREMENTS

Justification for, and Value of, a Data Dictionary

- o In examining the justification for DDs advanced by respondents, it is apparent that formal justifications, in terms of cost/benefit analysis, are generally missing. Only three respondents identified cost as a factor, as shown in Exhibit V-1.
- o 'Control' was advanced more times than any other single factor, although 'data' related factors as a group were more often mentioned. Justification is often done on a 'must have' basis rather than a formal evaluation: This is borne out by the number of DD users which developed their own independently of the availability of packages.
- o From a marketing viewpoint, these results imply that a DD should be sold as a tool to let a user know what data it has and be able to control it. The 'play' is on the user's fear of its shop getting out of control.
- o When respondents were asked to evaluate specific criteria which they may have used in their justification, all of them that replied stated they had considered reducing data redundancy in justifying a DD, as shown in Exhibit V-2. However, only one respondent had advanced that justification in so many words when asked to volunteer justification criteria.
- o Again the standards-related question drew a strong response with only two respondents replying negatively. Since it also had a relatively high proportion of voluntary support, standards establishment and enforcement related to data should be a strong sales emphasis.
- o Preparing for a DBMS came off poorly in this evaluation, probably because most respondents had a DBMS before getting a DD. However, as shown in Exhibit V-3, respondents overwhelmingly favor getting a DD before a DBMS. The reasons for the discrepancy are the availability of

EXHIBIT V-1

JUSTIFICATION CRITERIA FOR A DD
ADVANCED BY RESPONDENTS

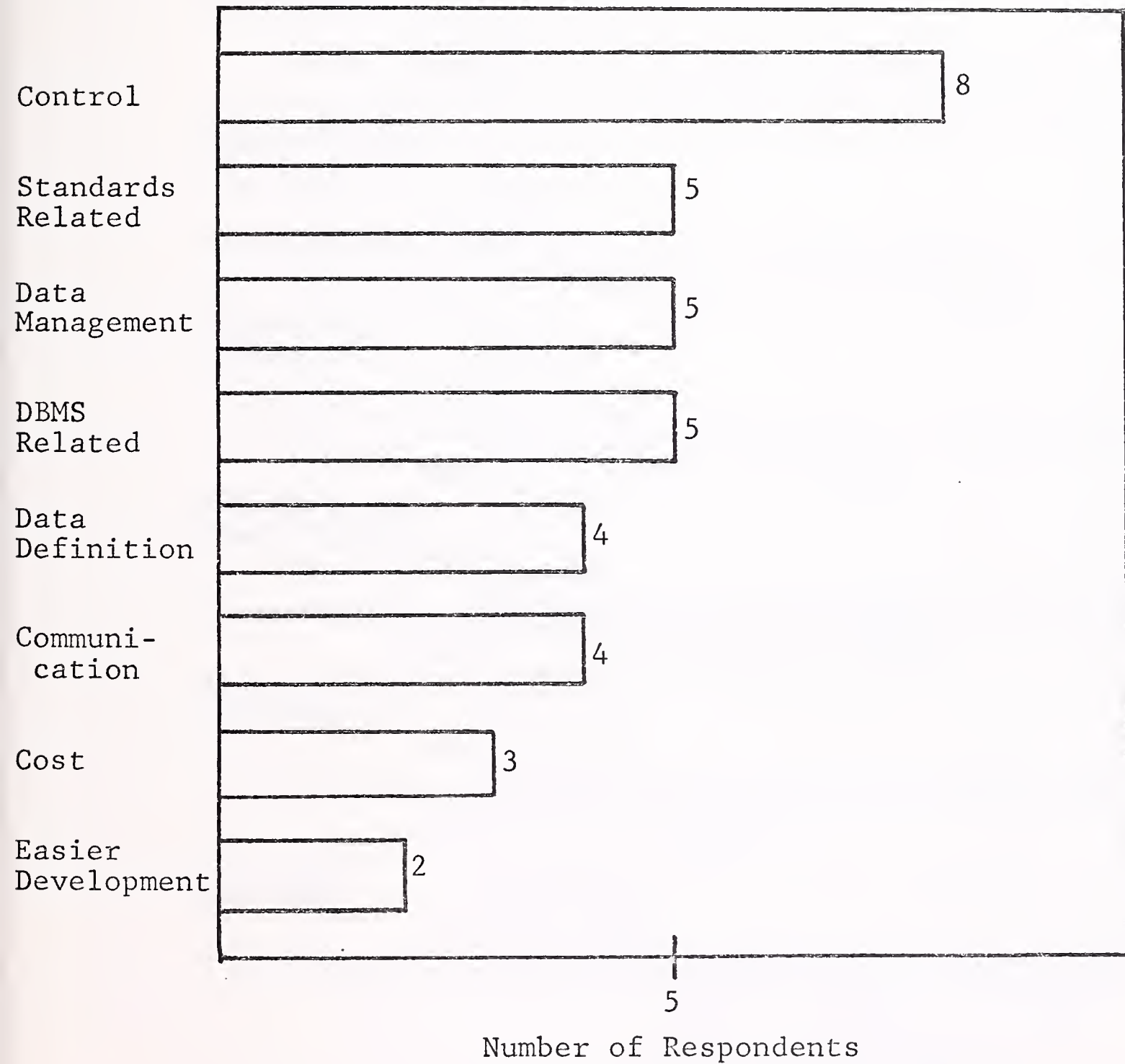


EXHIBIT V-2

RESPONDENTS EVALUATION OF SELECTED
CRITERIA FOR JUSTIFICATION

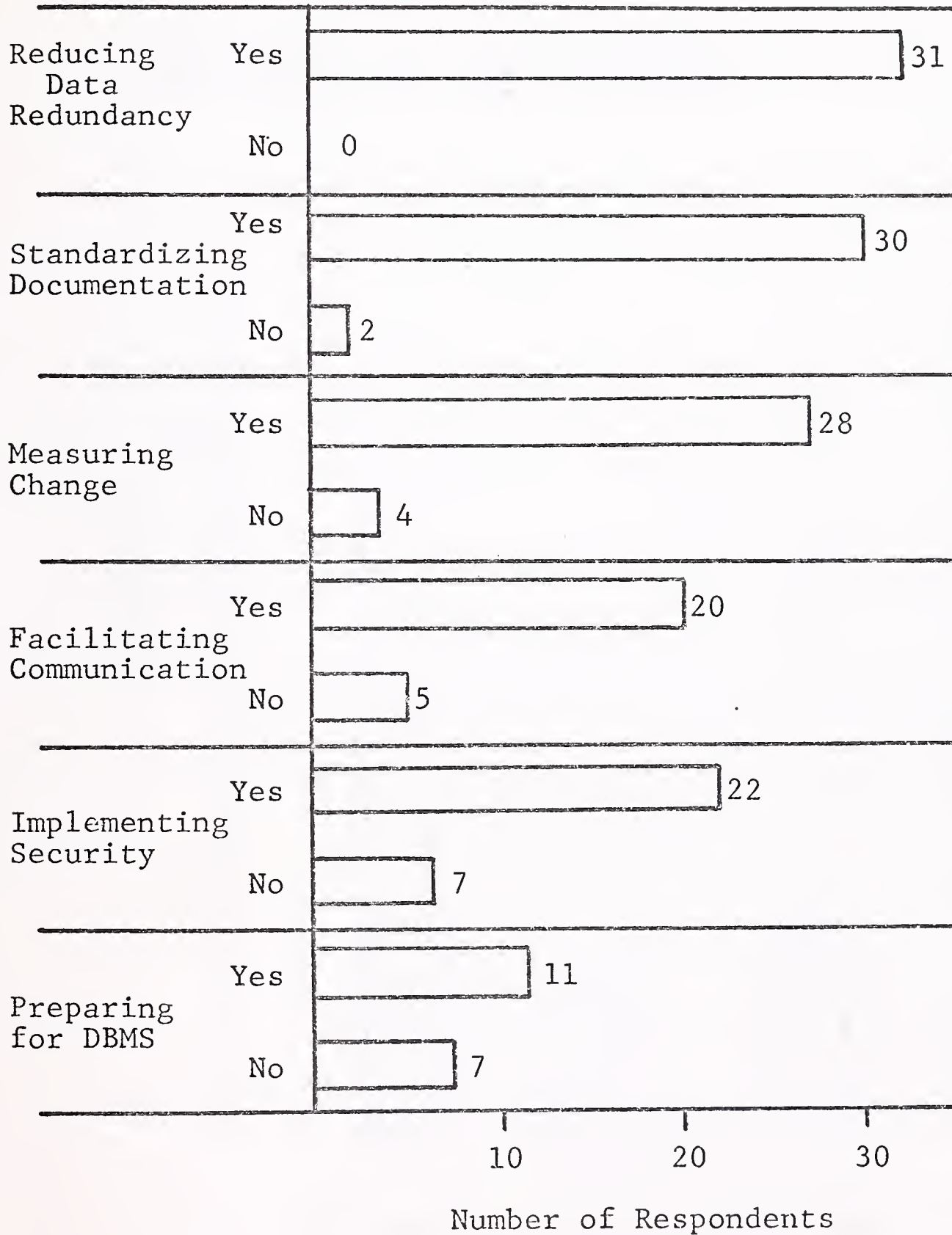
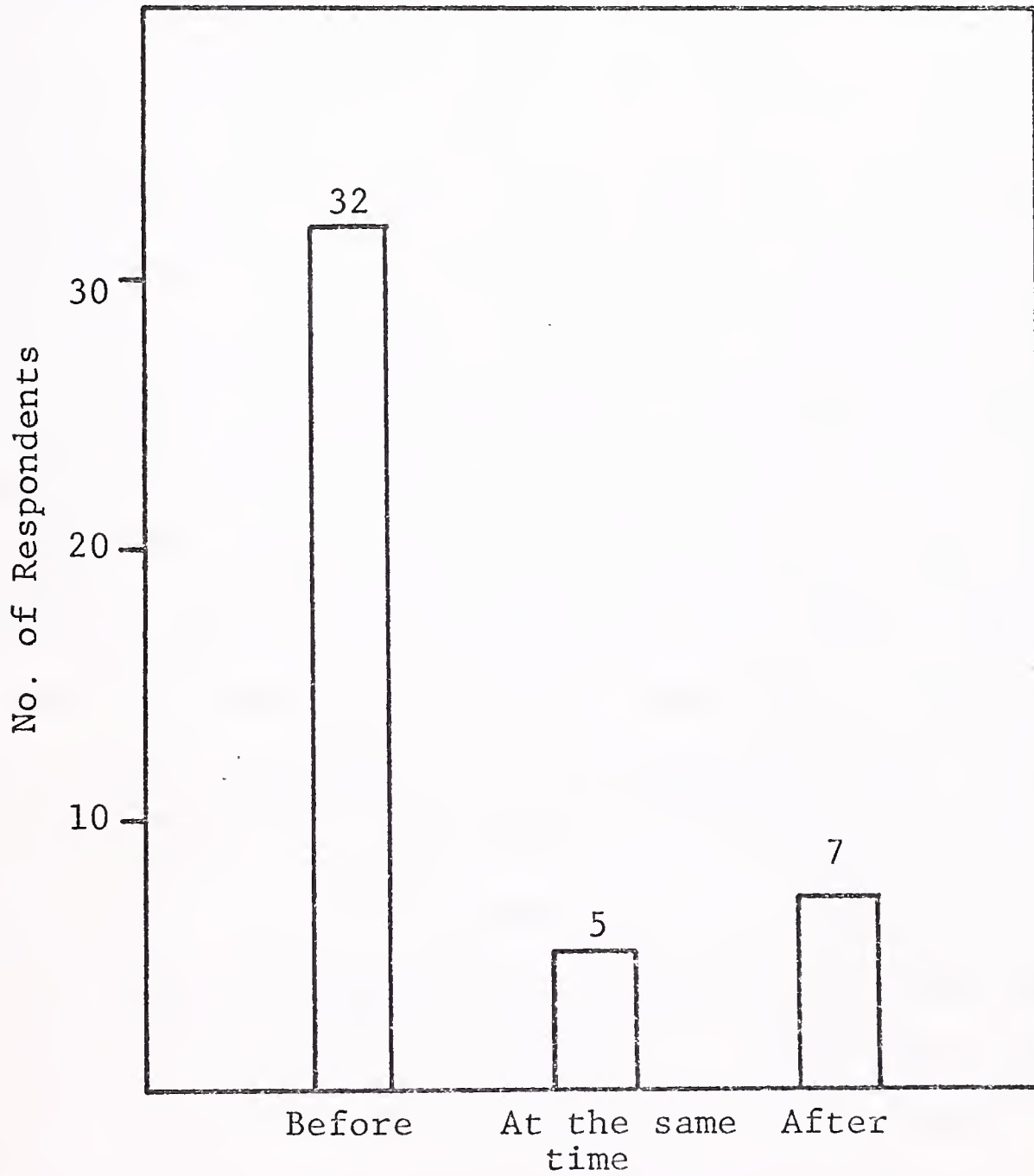


EXHIBIT V-3

RELATIONSHIP OF A DD TO A DBMS
FROM INSTALLATION DATE VIEWPOINT



DD BEFORE/AT THE SAME TIME/AFTER
A DBMS INSTALLATION

the product and the appreciation of its value.

Data dictionaries have lagged the market for DBMS. Many users have been forced to develop their own because of this lack of availability. Also, users of DBMS did not appreciate the value of a DD until after they had experienced the problems of using a DBMS such as IMS.

From a marketing viewpoint, DD package salespeople should emphasize the value of a DD in preparing for a DBMS. Targets should be those installations that will have DBMS over the next three years but have not yet got one. The sales pitch should emphasize the experience of others and their preference for getting the DD before the DBMS, had they had the opportunity.

o In examining closer the responses to the question on the relationship of a DD to a DBMS, it is apparent that several of the replies indicating that a DD should come after a DBMS were due to the fact that the respondent already had a DBMS, and therefore the DD had to come after, or that the only DDs that met their needs (UCC10 and IBM) required the use of a DBMS.

o None of the respondents contacted had any quantification of the value of a DD. One respondent likened the question to 'How valuable is your car to you?' The answer is unquantifiable; you simply have to have one. Some of the answers given are shown in Exhibit V-4: basically they repeat the justification criteria. This occurred throughout the interviews: respondents could not separate the justification criteria, selection criteria, and value of a chosen product.

EXHIBIT V-4

RESPONDENTS EVALUATION OF THE
REAL VALUE OF THEIR DD

Size of User	DD Used	Comment on Value
Large	Own	Saves time, effort and money.
Very Large	IBM	Better Management
Very Large	IBM	Enhanced program development, change control, and documentation.
Very Large	Data Catalog	Availability of data, readability of reports.
Very Large	Own	Documentation and control
Very Large	UCC 10	Generates IMS control blocks-Documentation
Very Large	UCC 10	IMS interface identified data. Control block generation.
Very Large	Lexicon	Reduced redundancy. Standards for compiler reference files
Large	Own	Control of data elements.
Very Large	Own	Program maintenance.
Very Large	IBM	Minimal, due to lack of use.

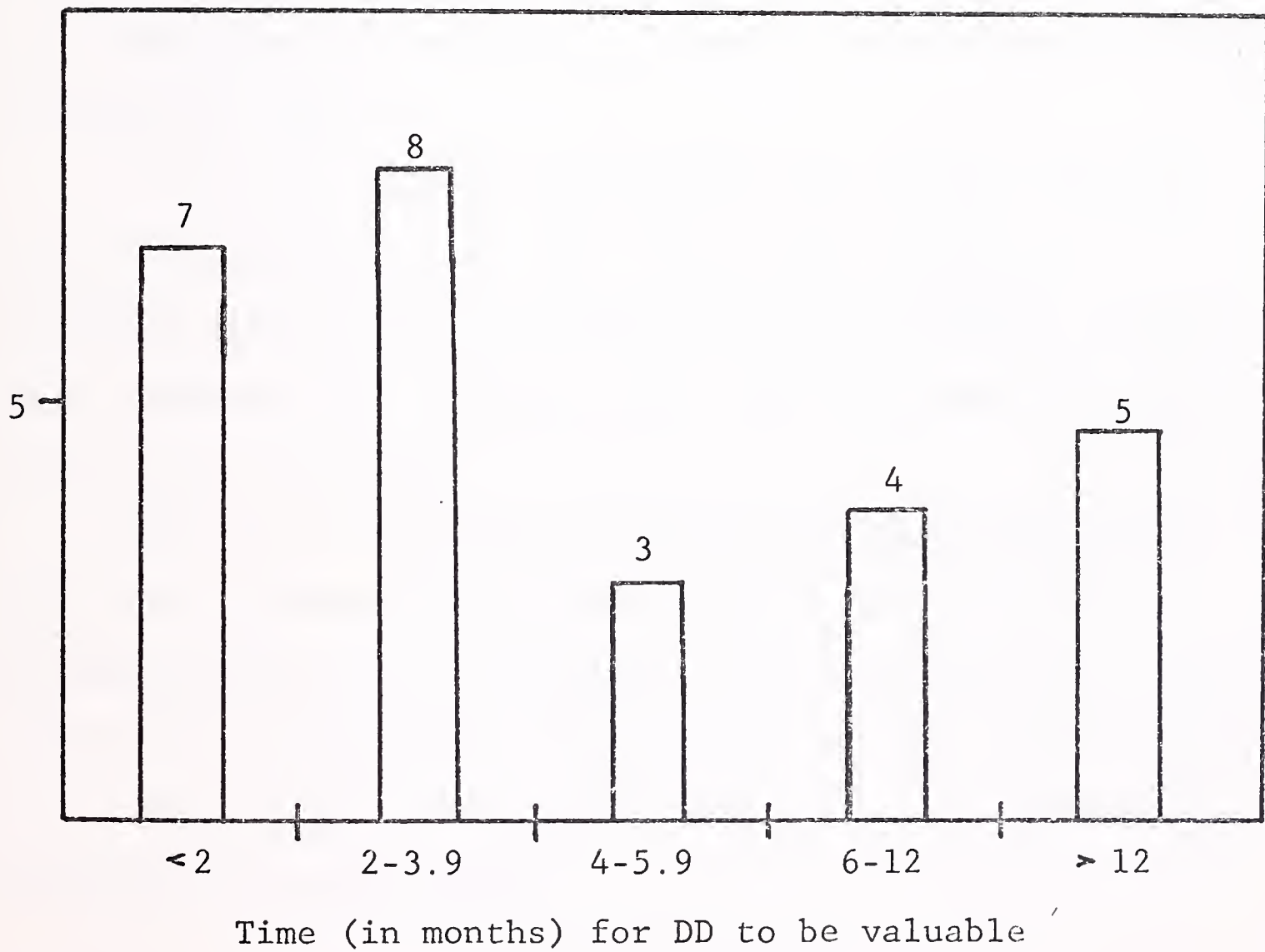
o The length of time it takes for a DD to be useful to an installation varies widely as shown in Exhibit V-5. From the results it appears that users who employ a DD to support development effort can obtain an immediate benefit; users which want to convert existing systems, find that it can take a year or more for benefits to be realized. All agree that the benefits are there and two thirds of the respondents considered their DD to be of real value within 6 months.

In marketing, therefore, the emphasis should be on the immediate benefit to be gained when a DD is used in development. It also means that the package should emphasize development support capabilities.

o In evaluating the number of users of a DD, there were two distinct groups present. The first group has six members and has an average number of users of six: the second group has three members with an average number of users of 70. The difference appears to lie in whether or not everyone has access to the DD, or whether only a select group, such as DBA staff or systems staff, has access.

EXHIBIT V-5

TIME TAKEN FOR A DD TO
BE OF REAL VALUE



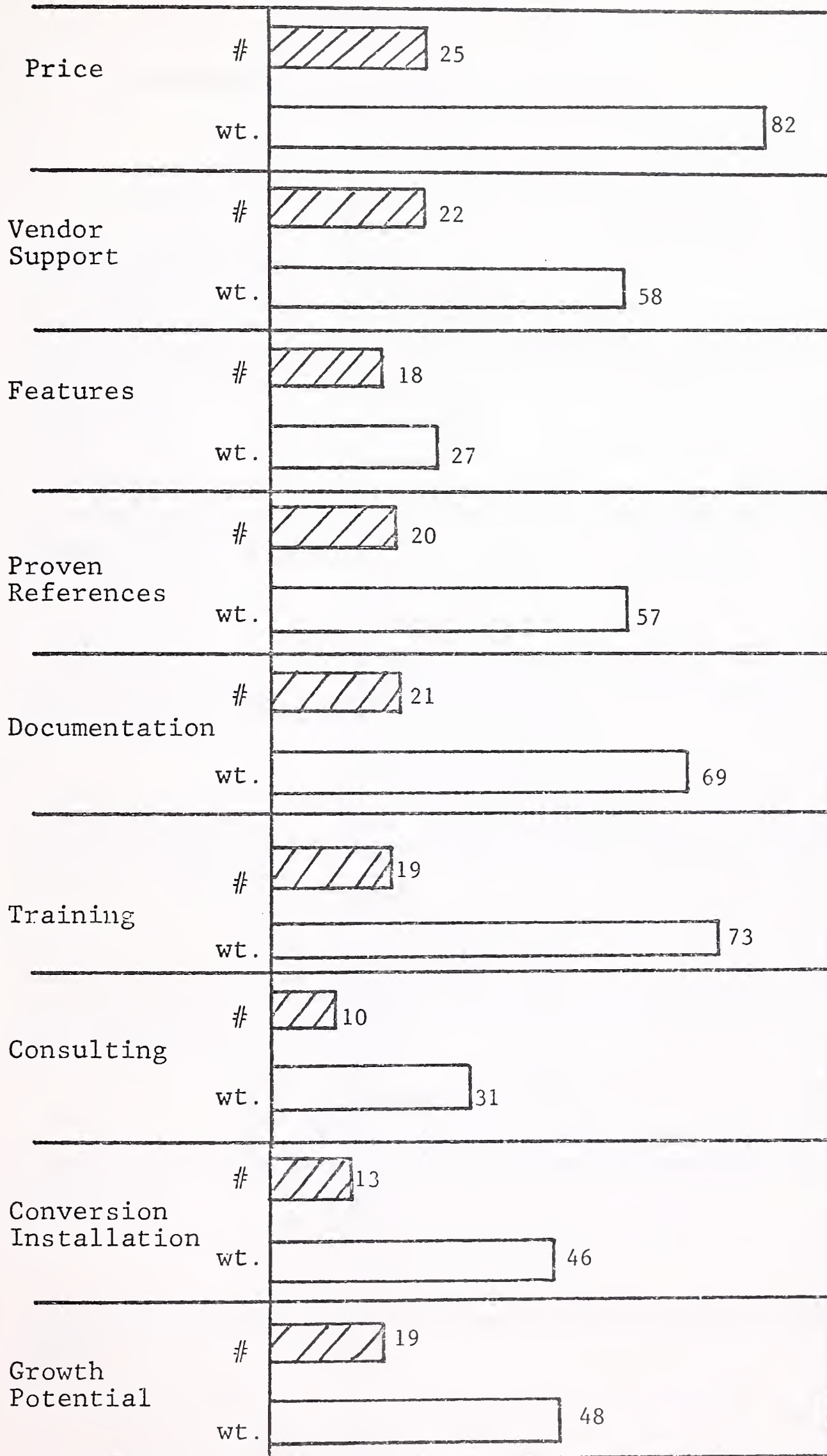
Selection Process and Purchase Level

- o When asked to rank a set of criteria used in selecting a package, respondents strongly favored 'Features', as shown in Exhibit V-6. On a scale where if every respondent ranked a feature 'Number 1' a perfect score would be 1.0, 'Features' achieved a 1.5 score.
- o 'Vendor Support' and 'Growth Potential' were ranked almost identically in second place, followed by 'Proven References'. 'Price' was simply not a major factor, tying with 'Documentation' in sixth place after 'Consulting'. 'Consulting' received fewer mentions than any other and should be regarded as of little importance for very large users. For large users it may be an advantage to have it.
- o The Data Base Administration department was the most frequently mentioned selection authority for a DD, as shown in Exhibit V-7. Purchasing authority usually resided with the top DP Manager, although Technical Managers such 'Manager of Application Development' were used more than 20% of the time.

The results of the research demonstrate that purchasing a DD is definitely a two-stage process with a technical selection preceding purchase authorization by a DP Manager. Then from a marketing viewpoint contact must be made at both levels: an initial contact through DP management probably being the best approach. The interest of both DP and DB management is obviously present.

EXHIBIT V-6
RESPONDENT EVALUATION OF PACKAGE SELECTION
CRITERIA FOR A DD

Weight



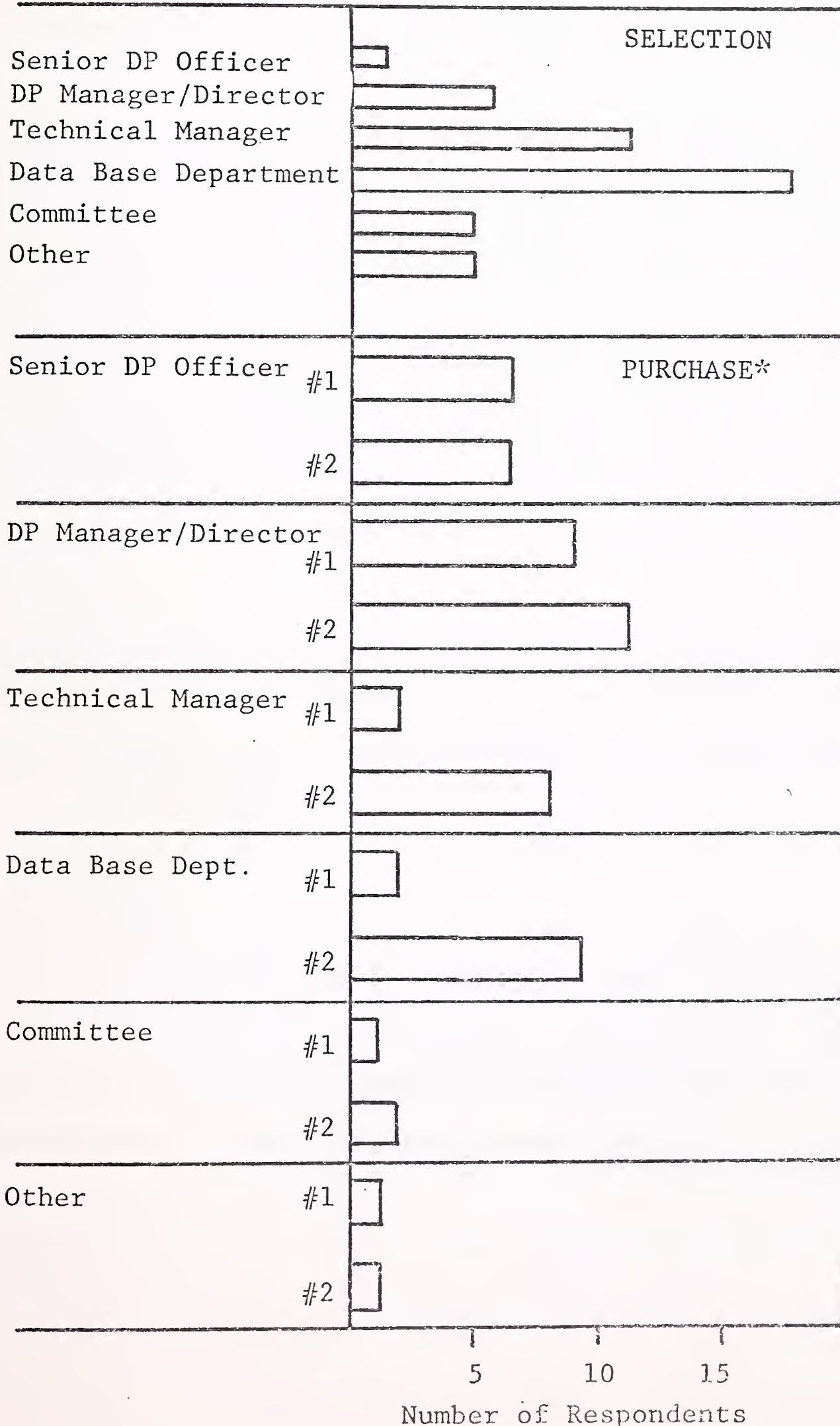
Weight based on 1=most important
7=least important

No. of Respondents per question 10 20 30 40 50 60 70 80 90
 Weight/Number of Respondents

INPUT

EXHIBIT V-7

SELECTION AND PURCHASING AUTHORITY
FOR A DD



*Two questions asked of respondents:
 #1 Asked of all early in interview
 #2 Only asked of those with, or getting a DD.

INPUT

Pricing

- o According to the survey, a DD is not a price sensitive product. Pricing ranked low as a selection criterion. Several respondents stated that the price of the package was trivial compared with the cost of operating a DD function.
- o In terms of their expected price to pay for a DD, Exhibit V-8 demonstrates the responses obtained. One respondent stated it would expect to pay \$20K for one installation and up to \$60K for use at all State agencies.
- o Although not price sensitive, several users commented that they did not want to pay for features they would not use. This indicates that options should be priced separately. Exhibit V-9 shows a suggested pricing mechanism - whereby a user would be able to select those options desired.

In this way a medium/large user would be able to start with a cost of \$10K and add a DBMS module when he obtained one, such as TOTAL. The initial price would be \$10K. On the other hand an IMS user wanting on-line and project management capability would pay \$25K.

Support Requirements

- o According to the respondents, little or no support is required, provided that the package operates according to specifications. It does not appear to be a technically complicated process and many of the respondents stated that they were virtually self-sufficient.

EXHIBIT V-8

EXPECTED PRICE OF A DD PACKAGE

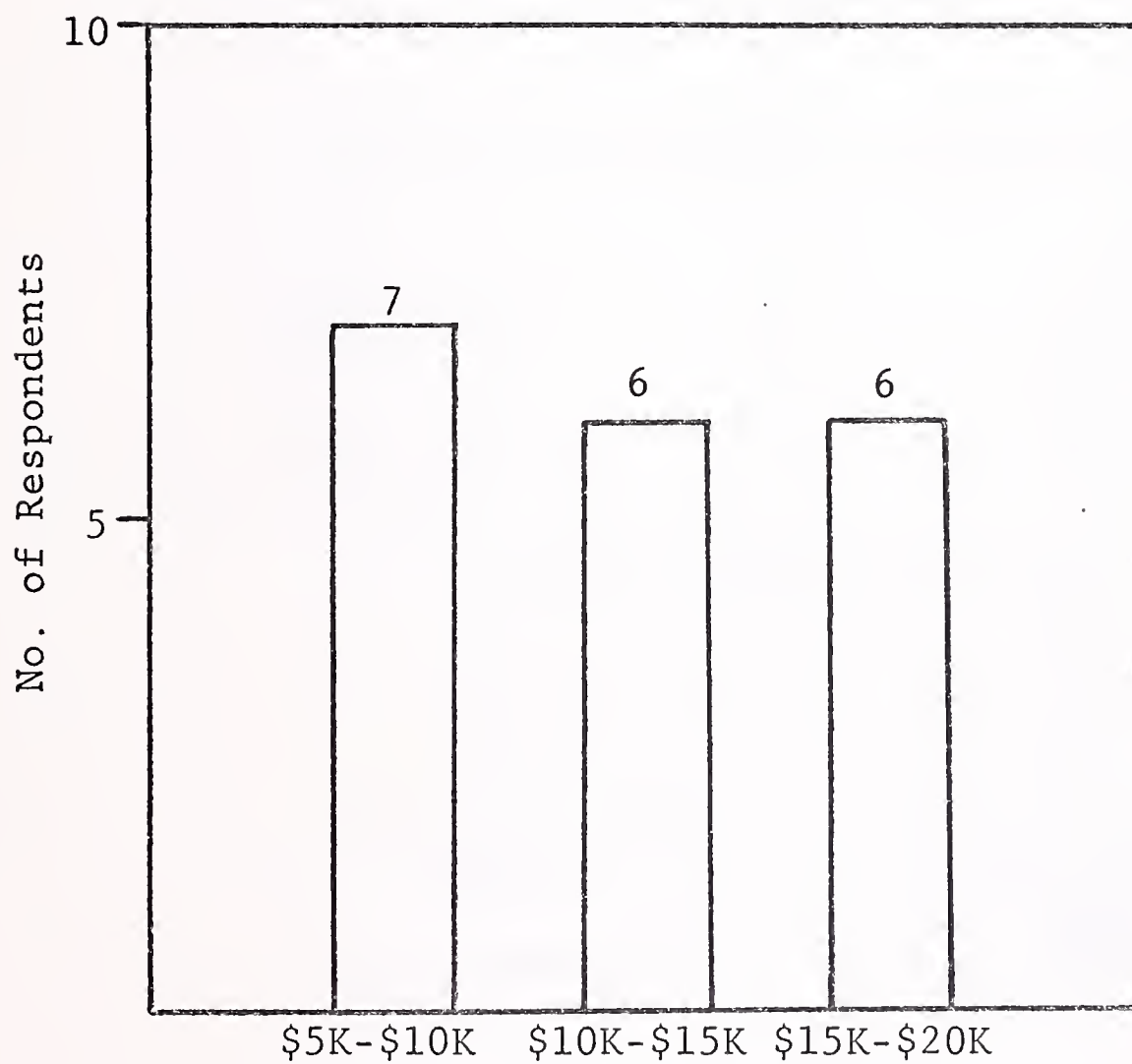
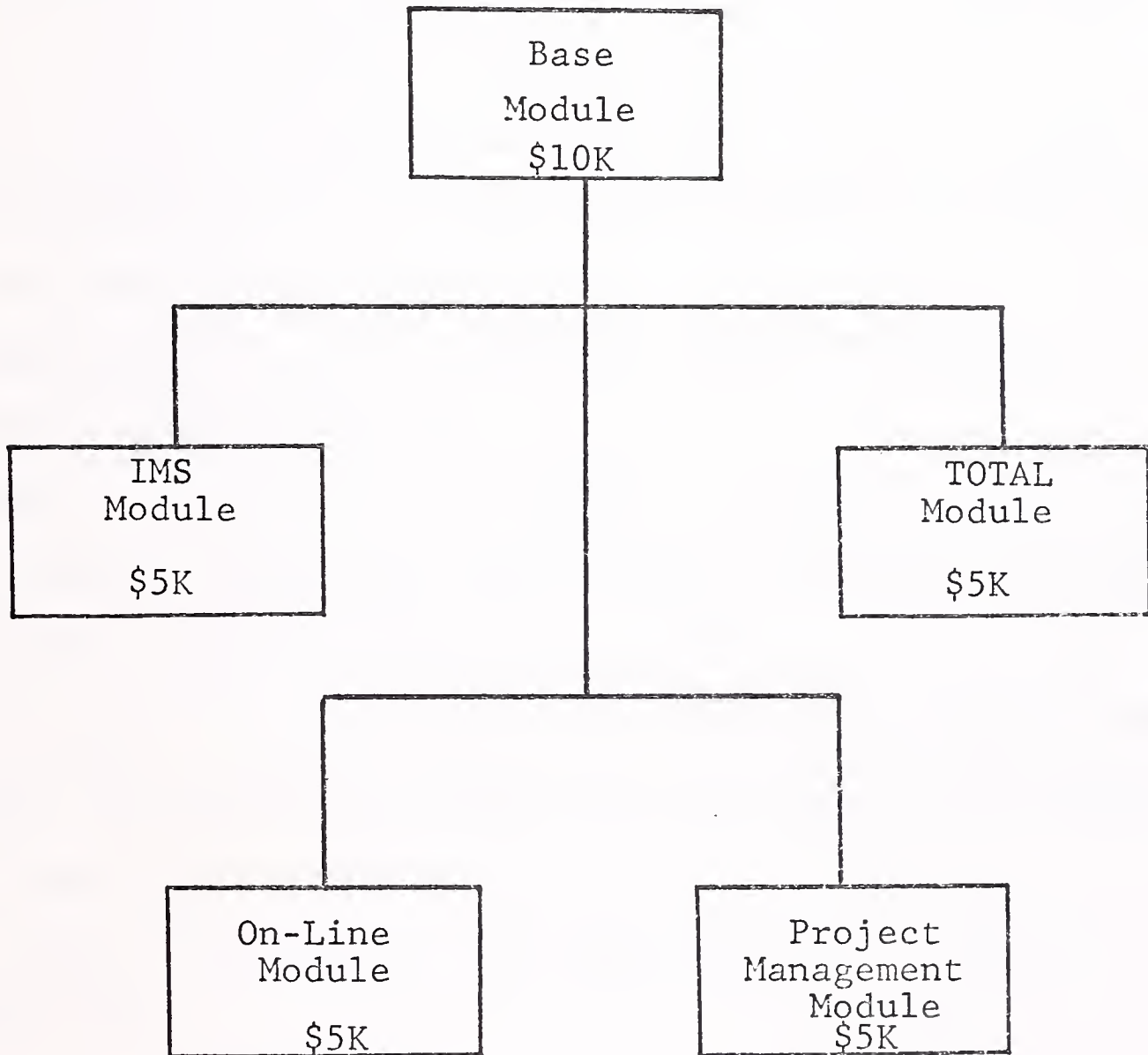


EXHIBIT V-9
SUGGESTED PRICING CHART



o The rank order of support requirements was:

1. Training
2. Education
3. Consulting/Conversion
4. Installation/Technical Support

Training was generally expected to occupy one or two days.

o One point that was raised in this context by several respondents was that it would be a distinct advantage to have an executive-level briefing. This would bring home to the corporate and top DP managers that the DD represented a new way for the DP department to do its business, rather than just another software package.

This ties in with several respondents comments on data being regarded as a corporate resource.

With all the problems of data security and privacy, as well as data control, corporate managers may well be receptive to such a presentation. This would be a powerful marketing tool, as well.

o All except two of the respondents who answered were satisfied with the support they received. A Lexicon and a UCC10 user both considered that the support could be improved.

o In terms of documentation, the required documents in order of importance were:

- User Manual
- Systems Manual
- Installation Guide
- Training Manual

o Overall the present quality of documentation does not appear good. Of UCC10, IBM's DD, Lexicon, and Data Catalog, only the last had no negative comments.

Since users place a lot of emphasis on useability of documents and there is a trade-off between support requirements and documentation as shown in the selection criteria, good documentation which is easy to read and use should be mandatory. The documentation should be separated for use at multiple levels of expertise and experience with DP.

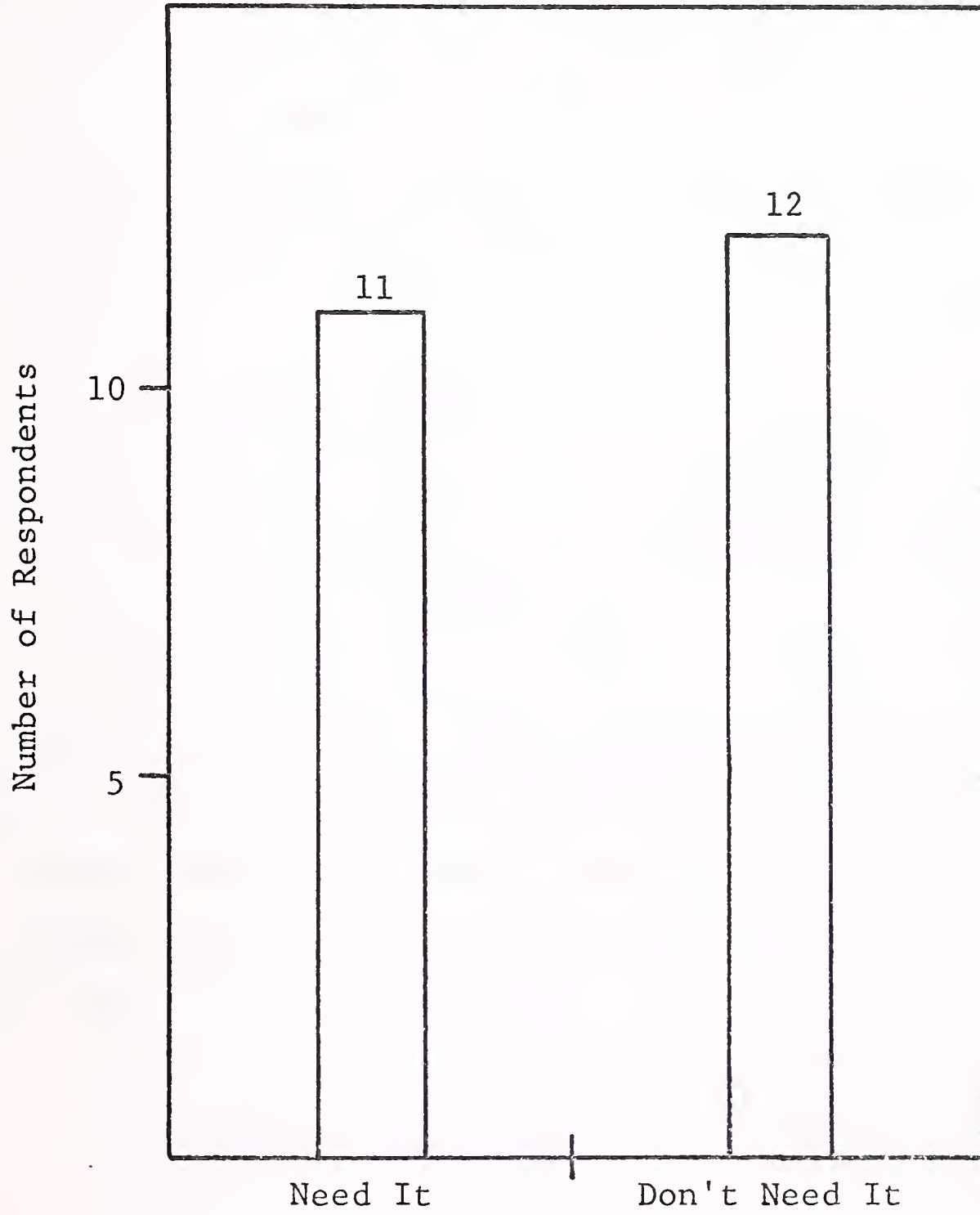
o Almost one third of respondents did require some on-site consulting to be available in order for them to install and use a DD. This was consulting for which they would be prepared to pay in addition to the package cost. Of the eight respondents which wanted it available, three were large users and the rest very large.

o The question as to whether or not users wanted source code resulted in an almost even split of respondents as shown in Exhibit V-10. Again this is something for which they would be prepared to pay extra. The source code should therefore be available.

As far as modification is concerned, the DD should provide 'user code' exits at appropriate places so that the user can add its own code without destroying the integrity of the package.

EXHIBIT V-10

RESPONDENTS' NEED FOR DD SOURCE CODE



VI COMPETITION

Respondents' Concerns with Data Dictionary Vendors

- o Respondents considered UCC 10 and IBM far more frequently than all other DDs, as shown in Exhibit VI-1. This emphasis was undoubtedly due to the high incidence of use of IMS in the sample.
- o Of the two instances where IMS users selected other than UCC 10 or IBM, Data Catalog was chosen. In these cases, UCC 10 was ranked high both times whereas IBM's DD was classified as

- 'Junk'
- 'Not a Class 1 product'

The reasons UCC 10 was not selected in these two cases were:

- 'Needs to address more than just DBMS data; it needs to attend to all corporate data'
- 'UCC 10 would not go under DOS'

- o Some comments on DDs and their vendors are shown in Exhibit VI-2. In spite of the negative comments, IBM has attracted and will continue to achieve sales based on future promise: five of the respondents selected IBM (over UCC 10 in most cases) because of their expectations. IBM's weaknesses appear to be:

- Reports and reporting
- Input handling and complexity
- Support

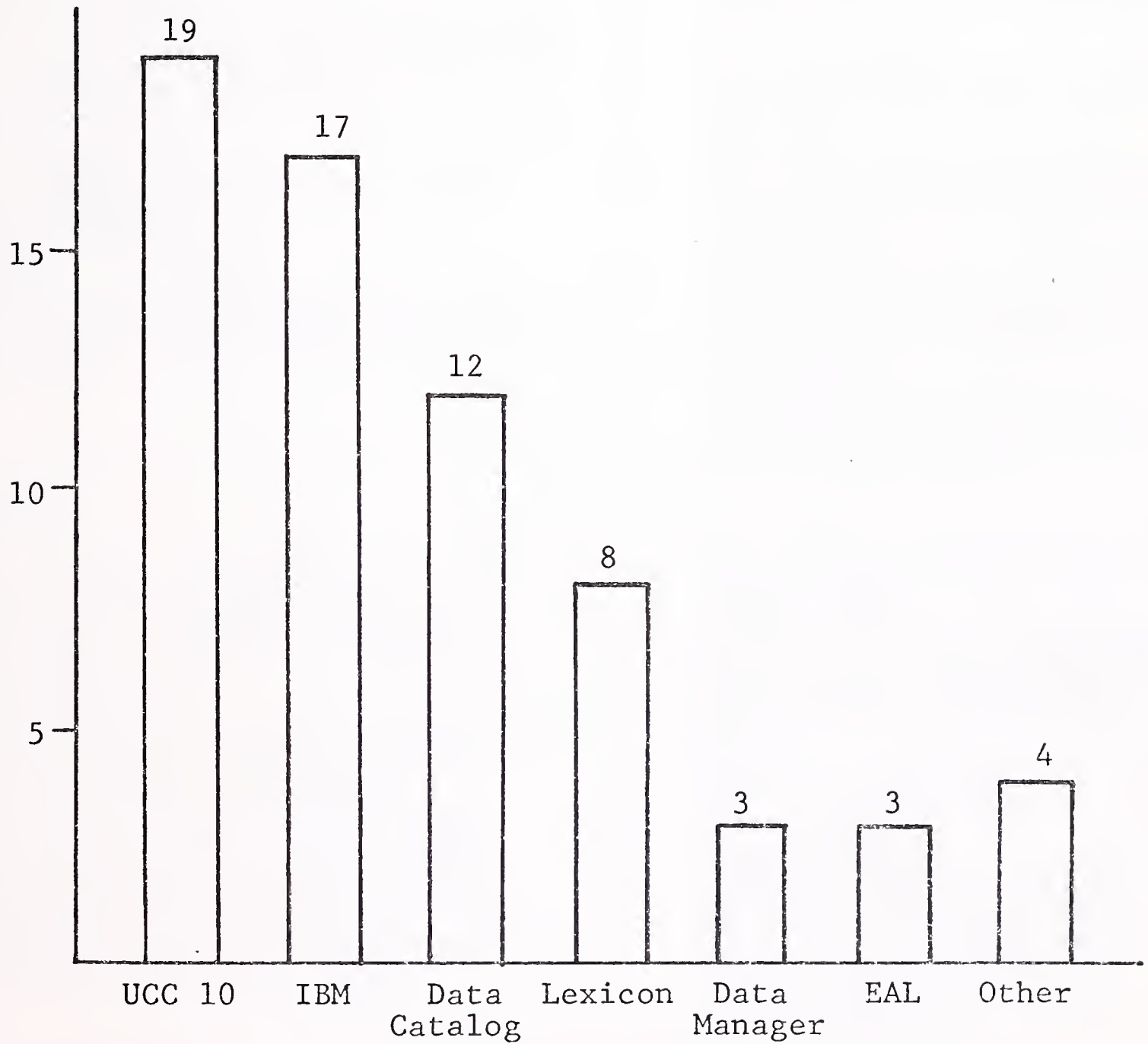
Strengths are:

- Future integration with IMS
- Interfaces with design aids
- Data structure and flexibility

- o UCC 10 is accepted as a better product than IBM by most respondents who compared them. However certain negatives are apparent

EXHIBIT VI-1

DDs CONSIDERED BY RESPONDENTS



(Out of 30 Respondents which gave specific vendors)

EXHIBIT VI-2

RESPONDENTS COMMENTS ON VENDORS AND DDs

IBM

- | | |
|---|--|
| <ul style="list-style-type: none"> o Selected on price(over UCC 10) now would be the same o Better off in future than with UCC 10 -DD structure best - reports lousy o Probably go with IBM (even though UCC 10 is better) expecting them to integrate into IMS o Didn't market well - not a class 1 product o Not adequate control-too much manual input o Not helpful people o 'Junk' o Dedicated to IBM-DBDs/PSBs more easily done with IBM(than Data Catalog) | <ul style="list-style-type: none"> o Dissatisfied-might switch to UCC 10 o IBM will develop the standard o Product not as responsive (as UCC 10) o Future features and tie to IMS design aids (over UCC 10) o IBM local support shaky o DB structure great-future support and development promised o Rejected because reports were not readable-By now DP people also had to have IMS o Product flexibility and user enhancement |
|---|--|

UCC 10

- | | |
|---|---|
| <ul style="list-style-type: none"> o UCC 10 more versatile than IBM lots of references -IBM none o Good job marketing. Price kind of high. No DOS so selected Data Catalog o Not adequate control - too much manual input - selected own o More features (than IBM) o Nice, if all you have is IMS o Rejected because needed IMS o Did not have reports for non-DP people and required IMS o Elaborate,complete, better than anything else available. | <ul style="list-style-type: none"> o UCC 10 more functions (than IBM) o Input done by clerical function chosen over IBM o Functional capability/ease of use-compared to IBM o Good testimonials o Data base structure not upgradable o Vendor stability would be -VE (2 respondents) o Rejected since has features not needed for which you pay. |
|---|---|

Data Catalog

- | | |
|---|---|
| <ul style="list-style-type: none"> o Only one which works with Univac 1108 o Seems OK- Must be stand alone o DOS and OS - Also input good - IMS user o Interfaces with TOTAL o Needs more than DBMS-TOTAL user | <ul style="list-style-type: none"> o UCC 10 had more functions IMS User o Didn't meet requirements IDMS User o Reports OK for non-DP interface with IMS if required. |
|---|---|

continued....

INPUT

EXHIBIT VI-2 continued

RESPONDENTS COMMENTS ON VENDORS AND DDs

LEXICON

- | | |
|--|---------------------------------------|
| o Presentation good | o Not IMS support enough |
| o Well done presentation -
wouldn't pay fee | o Selected since had
IMS interface |
| o Did good job - more for
their clients. | o Didn't want consulting |

OTHER

- | | |
|---|--------------------------------------|
| o EAL too simplified - no
control | o Pride-Logic - because
had Pride |
| o Data Manager - Main interest
because of ADABAS | |

including vendor stability and (apparently) not an equal commitment to future development and support. UCC 10 wins out continually on the 'features' side, at least as far as IMS is concerned. It cannot be used outside IMS and this caused it to be rejected in several shops which chose Data Catalog.

- o Data Catalog is the most flexible of the DDs on the market: it will interface with TOTAL and IMS, although not 'adequately' in many cases; it will work on Univac and Honeywell equipment as well as IBM; it can be used by non-DP people. Another strength of Data Catalog is its ability to take IMS and COBOL input to automatically generate DD entries. Finally it is simple to use.

- o Lexicon has been well presented by Arthur Anderson. However, it is cumbersome to use and many users don't want to pay the consulting fee that goes with it.

- o Of other packages, the only interest in DATA MANAGER stems from its relationship with ADABAS. Data Catalog will beat it out in IMS and TOTAL installation because of its interfaces, which DATA MANAGER doesn't have. Also Data Catalog is getting an on-line module which DATA MANAGER doesn't have.

- o Several TOTAL users stated they were 'Waiting for CINCOM' - however, if Eastern Airlines is the prototype for their DD, it will not be much competition for Data Catalog.

- o An important factor in selection of a DD, is knowledge of product. Most respondents appear to have considered all available vendors, with UCC, IBM and Synergetics achieving the most coverage. The Arthur Anderson presentations appear surprisingly good, and their coverage is extensive: several respondents which had not seen any other vendors had had an AA presentation on Lexicon.

Competitive Developments in Data Dictionary Market

- o IBM has two products available in the Data Dictionary market:
 - IMS Dictionary System. This is a Field Developed Program from the IMS Productivity Center in Palo Alto. This initially lacked capability, particularly in the cross-reference area: it has, however, been updated and IBM feels it is competitive with UCC 10 now. In fact, one salesman mentioned that 'a UCC 10 user was considering changing to IBM'.

This system will probably be integrated into IMS within the next two years which will make it tougher competition for UCC 10. Instead of being bought as a stand alone product it will be an option under IMS. The charge for the option will probably be slightly higher than for the Field Developed Program: users of the latter will be able to update to the IMS option for a one time charge.

This is categorized as an IMS aid; there are 12 of them at the moment including DB MAP and DP PROTOTYPE.

- The second IBM DD is IBM's DD/D with Data Element Glossary, which is an Installed User Program. It operates outside any DBMS. The current version is a new release of the old product or a derivative of it. In any event, this does not appear to be strong competition to the generalized data dictionary market right now.

- o HOSKYNS (now a subsidiary of Martin Marietta Data Systems) has a package known as HMS which has some attributes of a DD. It is most often used with Hoskyns' Manufacturing Application System (MAS): its main function is to allow rapid modification of this standard application software to individual clients' needs.

HMS consists of four modules:

Program File Processor - generates multiple program skeletons

INPUT

Record Data Processor - generates multiple record
layouts

COBOL Table Master - preprocesses program tables

COBOL Test Master - tests COBOL subprograms

o The PFP and RDP sell for about \$65k. They are used by Hoskyns staff in their project management consulting assignments as well as in MAS. This is analagous to Arthur Anderson's use of Lexicon.

o Lexicon now has about 45 installations. Arthur Anderson has just released a DOS/VS version. Lexicon will be a competitor to any generalized data dictionary, particularly when an organization is going to implement a data management or data base administration function at the same time, since AA provides the package as part of their setting up such a function on a consulting basis.

However, of their 45 installations probably only 30-35 use it to any extent and of these that do most of them have AA staff working on a consulting assignment. It is cumbersome to use and lacks the features of a Data Catalog.

Arthur Anderson is working with CALTEX, United Airlines, Houston Power and Light and others on interfaces with IMS, TOTAL, and other DBMS.

One positive feature is that they already have an interface with PANVALET.

o PRIDE-Logik is tied completely to the use of PRIDE which is an approach of M. Bryce and Associates to system design. Apparently there are over 450 installations of PRIDE including, GM, GE, RCA, BENDIX and Blue Cross/Blue Shield.

Logik is shipped to existing PRIDE users. No on-site installation help is needed or provided.

Basically it organized information for data base management but doesn't interface with any particular data base. Software AG has contacted them to work with ADABAS.

They will not provide source code. Do provide a dictionary of some 4000 attributes from which users may select for data element description. Users can search an attribute in any order required - but no 'AND/OR' logic is included.

It can be put on-line through ROSCOE, but Bryce doesn't think that's justified.

This will be a competitor to generalized data dictionaries in PRIDE installations. However, it and Lexicon require commitments many users may not feel they can make.

o Eastern Airlines DD is not a competitor except in relationship to TOTAL installations. It does not have acceptable capabilities and if CINCOM's product is based on this, it will not be a strong competitor to Data Catalog.

o Data Catalog is the one to beat for generalized data dictionaries. It has an on-line function just being released. Among its advantages are:

- Operates on IBM, Honeywell, and Univac
- Interfaces with IMS and TOTAL
- Works from COBOL statement input
- Works with data outside computer environment
- Easy to use

Its disadvantages include:

- Not working with PL/1
- Lacks features
- Interfaces with IMS and TOTAL are not automatic
- Lacks large user base

Data Catalog has only 50 installations now but it appears to be starting to move: it is the prime non-specialized data dictionary. It needs improvements but these are apparently in process.

o DATA MANAGER will not be competitive in installations requiring any kind of DBMS interface since getting a DD is most often linked to the acquisition of a DBMS this is a most serious drawback.

