A cluster evaluation of Navy quality of life programs

Michael J. Schwerin[a,*], Paul G. Michael[b], Dale N. Glaser[c], Kara L. Farrar[b]

[a]*Navy Personnel Research, Studies, and Technology, Navy Personnel Command, Institute for Organizational Assessments,*
5720 Integrity Drive, Millington, TN 38055-1400, USA
[b]*University of Memphis, Memphis, TN 38152, USA*
[c]*Pacific Science and Engineering Group, San Diego, CA 92122-5918, USA*

Abstract

The objective of this study was to describe a program evaluation system that can be applied to a wide range of quality of life (QOL) programs in the Navy—specific enough to capture the variability of several QOL programs with sufficient consistency across programs to facilitate multi-program comparisons. Evaluation measures were designed specifically for each program consisting of a core set of items that evaluated how well each program affected program-specific outcomes, general QOL program outcomes, and organizational outcomes (i.e. overall QOL, readiness, and retention plans). Surveys were administered at 13 QOL programs in three US Navy Fleet concentration areas: Southern California, Norfolk Virginia, and Yokosuka Japan. Program evaluation surveys were distributed to program patrons at the point of service for a 6-month period (n 10,765). Analyses and results examining general QOL program outcomes and organizational outcomes are presented. Advantages, limitations, implications, and areas for future research will be discussed. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Program evaluation; Cluster evaluation; Performance management; Outcome-impact assessment; Quality of life; Military and Navy

Cluster evaluation is an approach commonly used by national grant foundations (i.e. William K. Kellogg Foundation, Ann E. Casey, Public Welfare Foundation) to "evaluate a program that is being administered at different [autonomous] program sites aimed at bringing about a common general change" (Sanders, 1997, p. 397). The emphasis of the evaluation can be either formative (i.e. process-oriented) or summative (outcome-oriented), although most appear to be summative evaluations. The basic characteristics of a cluster evaluation are: "(a) it is holistic, (b) it is outcome-oriented, (c) it seeks generalizable learning, and (d) it involves frequent communications and collaborations among the partners" (Sanders, 1997, p. 399). The end-users of a cluster evaluation are the local program managers as well as the corporate-level program managers. The objective of this study is to describe the development of a quality of life (QOL) program evaluation system that can be applied to a wide range of QOL programs in the Navy—specific enough to capture the variability of several QOL programs with sufficient consistency across programs to facilitate multi-program comparisons.

At the time this evaluation was conceptualized, Navy QOL program were organized into several primary functional areas. Each primary functional area functioned autonomously but received funding and reported to a common, headquarters-level QOL program director. Much like many public and private sector agencies, Navy program area managers were being asked to defend their program's funding. Similar to program performance management approaches (e.g. Outcome—Asset Impact Models; Reed & Brown, 2001), Navy QOL program managers sought to demonstrate the effectiveness of their program by identifying the benefits and positive impacts that was a result of these programs. Toward that end, each functional area manager developed an approach to evaluate their program. Unfortunately, functional area managers used different approaches to evaluate their programs. One functional area evaluated their program with a random sample mail-out post-test only customer satisfaction survey. Another functional area employed a random sample mail-out post-test only need assessment survey. A third program area used customer satisfaction survey data to build a return on investment metric that represented the value of their program to the Navy. Other functional areas used either comment cards or had no customer data from which to

* Corresponding author. Tel.: +1-901-874-4654; fax: +1-901-874-2571.
E-mail address: michael.schwerin@persnet.navy.mil (M.J. Schwerin).

evaluate their programs. Upon receiving input from each of the functional area managers, the QOL program director was at a loss to prioritize programs for funding due to the variation in evaluation approaches.

Although each functional area manager used a reasonable approach to evaluate their program, the headquarters-level QOL program director for all of these QOL programs identified the need to develop one methodology that would facilitate comparisons among a broad range of programs. Additionally, rather than using customer satisfaction data to demonstrate program effectiveness, the QOL program-wide evaluation needed to employ outcome based measures. The emphasis would be on determining how well QOL programs met patron needs and what impact each program had on larger organizational outcomes, such as global QOL, readiness, and career intentions.

Kerce (1998) designed a program evaluation system that targeted a broad range of QOL programs in the United States Marine Corps (USMC). This research focused on an evaluation of programs based on a program's primary objectives or *reasons for being* (RFB). An RFB index, comprised of seven program objectives, was developed for use with the USMC. Preliminary work for the current study resulted in an expanded RFB index comprised of nine RFBs for the Navy (Schwerin, Michael, & Uriell, 2002b). Headquarters-level QOL program managers for the Navy then selected the RFBs that were most relevant for their program and these RFBs were included in the program evaluation measures. For example, the program manager for fitness and recreation programs selected RFBs that highlighted their primary objective of enhancing health, helping patrons concentrate on their jobs, and showing that the Navy is concerned about Sailors and their families.

This paper reports on the approach used by the US Navy to examine two types of summative outcomes—program outcomes (i.e. has the program met its primary objectives) and military outcomes (QOL, readiness, and retention plans) for 13 Navy QOL programs. Specifically, we will describe how evaluation measures were developed that would serve as a common measure for each of the 13 QOL programs, the research design and data collection procedure, analyses that were conducted, as well as conclusions and recommendations.

1. Cluster evaluation design

1.1. Focus of evaluation

The focus of this evaluation was to provide the headquarters-level QOL program director with information regarding the impact QOL programs had on meeting Sailor QOL needs as well as affecting overall readiness, and retention plans. Although the sponsor was interested in determining the number of QOL program users as well as reasons for use/non-use, their primary interest was with the perceived impact of QOL programs on Sailors and their families. The sponsor of these programs has responsibility for a wide range of programs. Programs included in this evaluation are the Child Development Program, the Food and Hospitality Program, the Recreation Program, the Youth Programs, the Clinical Counseling Program, the Deployment Support Program, the Personal Financial Management Program, the Relocation Assistance Program (SEAP), the Transition Assistance Management Program (TAMP), the Nutrition Education and Galley Services Program, the Chaplain religious enrichment delivery operation (CREDO), and the Navy college program (NCP).[1]

Some programs have high program use (e.g. fitness centers may serve over 300 patrons daily) while others have significantly fewer patrons (e.g. counseling services may serve three patrons a day). Another challenge of this study was to develop a way of evaluating programs that have seemingly unique objectives. For example, some programs focus on fitness, some on child care, some on family support but they all support patron QOL.

1.2. Instruments

Program evaluation measures were composed for each of the 13 programs included in this study. Each measure had a section that asked demographics (respondent demographics and program demographics), program-specific outcomes, RFB items, and organizational outcome questions.

The demographics had a subsection that focused on respondent demographics, asking program patrons for their social security number (SSN),[2] status in the military (active duty, reserve, retiree, family member, government civilian), and the date they completed the survey. Respondent SSN was asked in order to validate the self-report item measuring retention intent. Several studies have provided conflicting evidence as to the validity of a self-report measure of retention intent in the military (Cavin, 1987; Sharma, 1994). This study used the self-report measure of retention intent rather than their actual retention data. The main reason being that it typically takes between 6 and 12 months for active duty military personnel (i.e. officer and enlisted) to separate from the military and this time span would have not allowed researchers to link evaluation results to actual retention behavior.

Program demographics were asked to gather some

---

[1] A complete description of each program, evaluation measures, and additional results can be found in Schwerin et al. (2002b).

[2] Participants were informed of their rights by way of the Privacy Act Statement and informed consent information on the first page of the survey. Informed consent is presumed when the participant reads the informed consent statement and subsequently completes the survey. Completed surveys were sealed in return postage-paid envelopes and mailed to the authors for analysis. Data were reported at an aggregate level so that no individual could be identified. Program representatives were instructed that participants could chose not to report their SSN or could decline to participate in the study.

general information how patrons learned about the program and which components of the program they were using. More specifically, items focused on which component of the program they were using when they were asked to complete the survey (e.g. for Fitness and Recreation were you using the Fitness Center, Golf Course, Bowling Center, etc.), why they chose to use the program, and how frequently they used the program.

Program-specific outcomes were included to assist program managers and coordinators in learning more about how well services were provided. Several questions focused on aspects of program quality—quality of service, customer service, range of services, value for your dollar, and hours of operation. Other program-specific outcomes were related to the program objectives (e.g. Fitness and Recreation program asks patrons what effect this program has on unit morale and the sense of community).

RFB index items reflect the primary purpose(s) of Navy QOL programs. A set of seven RFBs were developed in previous work with the USMC (Kerce, 1998; Kerce, Sheposh, & Knapp, 1999) and expanded to an index of nine RFBs for this study. These RFBs reflect how a QOL program serves to:

1. promote the physical and psychological well being of members, maintaining QOL at a level to *attract qualified men and women to the USN*;
2. promote the physical and psychological well being of members, maintaining QOL at a level to *retain qualified men and women in the USN*;
3. provide a level of support that allows members to *concentrate on their mission*.
4. provide a level of support that allows *availability for deployment*;
5. provide educational opportunities in order to maintain Navy expertise;
6. demonstrate *concern for members and their families* to enhance morale and commitment to the USN;
7. help to ensure the health and safety of USN personnel and their families;
8. make available the skills and tools to *adjust to the stresses of military life*; and
9. increase personal and family satisfaction with the military life.

This RFB index represent the reason a program exists in the Navy. Although not necessarily independent of one another, RFBs were measured independently with separate survey items to represent unique program objectives. QOL program managers selected several of their primary RFBs from this list for inclusion on their program evaluation measure. Program managers may have selected the same RFBs for different programs since they may be serving QOL objectives common to each program (e.g. ensuring the health and safety of personnel and their families may be an objective of fitness and recreation as well as nutrition

education and galley services). Since this list of RFBs was constructed to reflect *any* reason the Navy might provide a QOL program, some RFBs may not have been selected for inclusion on a program evaluation measure. For example, the RFB 'attract qualified men and women to the Navy' was not selected for any of the programs currently being studied but it might have been an RFB for the Navy Music/Band program since its primary functions are ceremonial bands and recruiting.

Earlier QOL program evaluation efforts (Kerce, 1998; Kerce et al., 1999) have stopped short of the key organizational outcomes that are of particular interest to the Navy. Specifically, the Navy is keenly interested in not only how a program has affected QOL, but also a program's impact on QOL, readiness, and retention. Conceptually, the outcomes of QOL, readiness, and retention plans are somewhat related. QOL programs may improve global perceptions of QOL, therefore affecting readiness. Causal modeling of the potential mediating effect of QOL on the outcomes of readiness and career intentions show only a partial mediating effect for several programs (Schwerin, Michael, Glaser, & Uriell, 2002a). Although the relationship between QOL and readiness is statistically significant, its practical significance would not preclude these outcomes from being evaluated independently. The following military outcome measures have been developed for this program evaluation to measure how well a program:

1. contributes to my QOL in the military;
2. impacts my overall readiness (i.e. the extent to which I am prepared, able, and motivated to perform my job);
3. impacts my retention decision.

Additionally, a career intent question was included to determine the career plans of program users.

1.3. Design and procedure

The research design for this evaluation employed a post-test only using a purposive sampling approach (Levy & Lemeshow, 1999) where individuals were selected who were considered to be most representative of the population as a whole. Program evaluation measures were administered on-site to *experienced program users* over a 60-day period. Because QOL programs varied greatly in the number of patrons they served, a random sample mail-out survey, by chance, would fail to sample enough users of small, specialized programs to conduct meaningful statistical analyses. This is problematic since these specialized programs with few users meet significant QOL needs (e.g. counseling services, youth programs, financial management). The evaluation team decided to employ a targeted on-site data collection in order gather data from large and small programs to facilitate program comparisons.

Experienced program users were defined as patrons/clients who have used the program more than once or, in

their estimation, have a good familiarity with the program and services provided. The reason for imposing a participation screen for those receiving a survey was that many QOL programs have a large group of infrequent program contacts and a smaller group of knowledgeable program patrons. Military personnel and their families re-locate typically every 3 years. Oftentimes, people come to a program just to familiarize themselves with the program resources (e.g. staff, hours, computer resources, equipment, etc.). Program managers preferred that surveys be administered to patrons who have enough knowledge of the program to accurately evaluate the program. While a set of survey items could have been developed to identify experienced program users, program managers felt that their staff could easily differentiate infrequent program contacts and frequent program users.

Program assessment surveys were administered at various QOL programs coordinated by Morale, Welfare, and Recreation (MWR), Fleet and Family Support Centers (FFSC), NCP, and Chaplain Religious Enrichment Development Operation (CREDO). Additionally, a program evaluation of the Nutrition Education and Galley Services program was combined with an evaluation with its point of implementation—Navy Food Services (i.e. shore-based galleys). Some QOL programs were excluded at this phase of the study primarily due to the sponsor's desire to select 'spotlight' QOL programs and expand on some more challenging programs once the initial evaluation was completed. For example, Sexual Assault Victim Intervention (SAVI; a program that provides counseling and victim advocacy) or Navy Casualty Assistance (a program that is involved with administration of death benefits to the families of Sailors who die on active duty) present methodological challenges that do not exist with the set of 13 programs under evaluation.

Surveys were administered at three Navy fleet concentration areas: two in the continental United States (East Coast CONUS and West Coast CONUS) and one outside the continental United States (OCONUS). Each of the CONUS and OCONUS data collection locations were selected by headquarters-level program managers because they were the largest, most representative Fleet concentration areas in each region and, although not representative of the Navy, these three Navy regions collectively represent approximately 40% of the Navy active duty population.

Data collection was conducted during a 6-month period from November 1999 to May 2000. Members of the cluster evaluation team traveled to each base within the three Fleet concentration areas. During each of the site visits, evaluators delivered the evaluation supplies (i.e. survey forms, pencils, and postage-paid business reply envelopes), described the purpose, approach, and anticipated benefits of conducting the program evaluation study. When possible, evaluators held additional meetings with the service providers to describe the study and answer any questions at that time. Program administrators were then asked to

administer the two-page program evaluation survey to experienced program users over a 60-day period. Additionally, program managers were asked to administer surveys to users of all sub-programs within a program (e.g. all fitness and recreation activities subsumed within MWR Fitness and Recreation). The program administrator was asked to briefly describe the purpose of the survey to the program user, tell them how the information would be used, and ask them to complete the survey on-site and seal it in the postage-paid business reply envelope provided.

1.4. Data analysis

For each program, the frequencies, means, and standard deviations were reported for the demographics (respondent demographics and program demographics), program-specific outcomes, RFB items, and organizational outcome questions. Additionally, results are reported separately for the active duty Navy respondents and the other groups (i.e. spouses of active duty service members, retirees, etc.) who responded.

In the technical report that was delivered to the evaluation sponsor (Schwerin et al., 2002a), two sets of multi-variate analyses were presented for each program. The first focused on a path analysis model produced using structural equation modeling (SEM) that was conducted to examine the potential mediating effect of QOL separately for readiness and career intentions. Additionally, these analyses helped determine the strength of the relationship between a QOL program-meeting patron needs (through the RFBs) and the outcome measures. The second analysis consisted of a multiple regression analysis conducted to examine the relationship of individual study variables on the outcomes of QOL, Readiness, and Career Intentions. For the purpose of this article, only the results for the RFBs and outcome items will be reported. Detailed results for program-specific outcomes as well as multiple regression analyses and SEM can be found in Schwerin et al. (2002a).

2. Results

Data were collected from a total of 10,765 program users in the three Fleet concentration areas over a 6-month period.[1] Since QOL program managers selected two or more RFBs that they viewed as representative of their program's reason for being, the RFB items on the evaluation survey varied.

2.1. Reasons for being items

Demonstration of concern for members and families. Means for the *demonstration of concern for Sailors and their families* RFB item for all programs are presented in Table 1, ordered from the highest to lowest. All patron ratings for this item were over 4.0 (associated with the

verbal anchor of *agree*) for each of the 13 QOL programs evaluated. CREDO and Recreation were rated the highest, while Food and Hospitality and Nutrition Education and Galley Services were rated the lowest.

Satisfaction with military life. Means for the satisfaction with military life RFB item for all programs are presented in Table 2, ordered from the highest to lowest. Patron ratings exceeded 4.0 (associated with the verbal anchor for agree) for six of the 13 programs while ratings for five programs approached 4.0. One program, TAMP, was rated by patrons at 3.48, closer to the verbal anchor for neither agree nor disagree. CREDO and Recreation were rated the highest, while TAMP and Financial Management were rated the lowest.

Concentration on job/duties. Means for the concentration on my job/duties RFB item for all programs are presented in Table 3, ordered from the highest to lowest. Patron ratings exceeded 4.0 (the verbal anchor for agree) for three of the 13 programs while ratings for three programs approached 4.0. One program, TAMP, was rated by patrons at 3.44, closer to the verbal anchor for neither agree nor disagree. Child Development, CREDO, and Youth Programs were rated the highest, while TAMP and Financial Management were rated the lowest.

2.2. Organizational outcome items

Perceived impact on quality of life. Means for the *effect of program on my QOL* outcome item for all programs are presented in Table 4, ordered from the highest to lowest. Patron ratings exceeded 4.0 (associated with the verbal anchor for agree) for 12 of the 13 programs while ratings for one program approached 4.0. Recreation and CREDO were rated the highest, while TAMP and Nutrition Education and Galley Services were rated the lowest.

Readiness program measure. Means for the *effect of program on your overall readiness* outcome item for all programs are presented in Table 5, ordered from the highest to lowest. Patron ratings exceeded 4.0 (associated with the verbal anchor for agree) for five of the 13 programs while ratings for the seven remaining programs approached 4.0. CREDO was rated the highest, while Food and Hospitality and TAMP were rated the lowest.

2.3. Correlation between program RFB mean scores and outcome items

The mean score for all RFBs for each program was computed and correlated to the organizational outcome items (readiness, QOL, and retention intent, see Table 6). There was a significant correlation between the RFB mean score and overall readiness and QOL for all programs. Additionally, there was a significant, although somewhat lower, correlation between the RFB mean score and retention intent for all programs except SEAP. Programs with the highest positive correlation to retention intent were

Table 1
Program patron ratings of *demonstration of concern* item

Program	Active duty Navy program patron				Non-active duty Navy program patron			
	$\bar{X}$	SD	Percent (%) reporting *agree*	Percent (%) reporting *disagree*	$\bar{X}$	SD	Percent (%) reporting *agree*	Percent (%) reporting *disagree*
CREDO Program	4.77	0.51	98.0	0.7	4.79	0.48	97.3	0.3
Spouse Employment Assistance Program (SEAP)	4.52	0.74	93.1	3.4	4.50	0.70	92.3	1.2
Youth Programs	4.47	0.70	93.8	1.6	4.27	0.80	88.1	2.8
Recreation Program	4.46	0.71	92.6	1.6	4.60	0.67	94.3	1.2
NCP	4.41	0.77	90.8	2.5	4.26	0.81	82.7	2.1
Clinical Counseling Program	4.35	0.80	88.0	2.6	4.35	0.79	92.1	3.0
Child Development Program	4.33	0.89	85.8	5.2	4.40	0.81	89.2	3.7
Financial Management Program	4.23	0.74	86.5	1.4	4.36	0.68	86.9	1.6
Relocation Assistance Program	4.19	0.78	84.3	2.2	4.31	0.72	89.4	2.1
Nutrition Education and Galley Services Program	4.04	0.91	76.8	5.3	4.08	0.84	73.4	2.4
Deployment Support Program	4.01	0.78	76.2	2.3	4.39	0.74	94.1	4.5
Food and Hospitality Program	3.99	0.91	76.6	6.2	4.07	0.81	81.5	3.2
TAMP	–	–	–	–	–	–	–	–

Note. TAMP not included since this item was not selected as a *Reason for Being*. Mean scores range from 1 = *strongly disagree* to 5 = *strongly agree*. Program totals are ordered highest to lowest by active duty Navy mean score.

Table 2
Program patron ratings of *satisfaction with military life* item

Program	Active duty Navy program patron				Non-active duty Navy program patron			
	$\bar{X}$	SD	Percent (%) reporting *agree*	Percent (%) reporting *disagree*	$\bar{X}$	SD	Percent (%) reporting *agree*	Percent (%) reporting *disagree*
Youth Programs	4.28	0.84	86.0	3.2	4.18	0.80	83.1	2.8
Recreation Program	4.27	0.80	86.5	2.8	4.17	0.88	84.8	4.4
CREDO Program	4.25	0.88	78.6	3.2	4.34	0.80	79.5	1.5
NCP	4.25	0.86	81.6	3.3	3.94	0.88	66.5	1.9
Child Development Program	4.04	1.06	76.9	11.7	4.11	0.98	78.3	6.9
Spouse Employment Assistance Program (SEAP)	4.03	0.78	79.3	3.4	4.05	0.83	79.3	3.0
Food and Hospitality Program	3.97	0.92	75.9	6.3	3.99	0.84	77.1	3.9
Relocation Assistance Program	3.93	0.88	71.4	4.4	4.07	0.82	78.8	3.9
Clinical Counseling Program	3.83	0.97	65.5	7.6	3.74	1.06	58.9	9.5
Financial Management Program	3.73	0.86	63.1	6.8	4.07	0.78	64.8	6.2
Deployment Support Program	3.61	0.84	55.4	6.8	4.07	0.82	82.0	6.0
TAMP	3.46	0.96	51.1	13.8	3.72	0.97	58.6	7.0
Nutrition Education and Galley Services Program	–	–	–	–	–	–	–	–

Note. Nutrition Education and Galley Services not included since this item was not selected as a Reason for Being. Mean scores range from 1 = strongly disagree to 5 = strongly agree. Program totals are ordered highest to lowest by active duty Navy mean score.

Table 3
Program patron ratings of *concentrate on job and duties* item

Program	Active duty Navy program patron				Non-active duty Navy program patron			
	$\bar{X}$	SD	Percent (%) reporting *agree*	Percent (%) reporting *disagree*	$\bar{X}$	SD	Percent (%) reporting *agree*	Percent (%) reporting *disagree*
Child Development Program	4.27	0.83	85.6	3.7	4.30	0.70	86.6	0.0
CREDO Program	4.22	0.90	77.7	3.5	4.19	0.93	75.6	4.9
Youth Programs	4.21	0.85	82.8	3.2	3.89	1.05	66.6	11.1
Recreation Program	3.96	0.93	68.2	5.3	3.92	1.02	68.4	7.9
Clinical Counseling Program	3.83	0.87	66.3	5.0	3.67	1.01	58.3	8.4
Relocation Assistance Program	3.80	0.84	64.9	4.4	3.94	0.90	70.6	5.9
Food and Hospitality Program	3.71	1.00	59.2	9.1	3.78	0.89	64.9	4.9
Nutrition Education and Galley Services Program	3.68	1.01	58.9	9.7	3.88	0.87	66.6	5.1
NCP	3.66	0.93	54.5	8.4	3.68	0.95	57.2	7.2
Spouse Employment Assistance Program (SEAP)	3.62	0.90	55.1	10.3	3.91	0.70	72.7	0.0
Financial Management Program	3.57	0.92	55.2	10.0	3.66	1.13	56.0	10.7
Deployment Support Program	3.54	0.90	52.6	8.9	3.50	1.00	25.0	0.0
TAMP	3.45	0.90	48.6	12.3	3.10	0.83	24.4	22.0

Note. Mean scores range from 1 = strongly disagree to 5 = strongly agree. Program totals are ordered highest to lowest by active duty Navy mean score.

Table 4
Program patron rating of *program effect on QOL* item

Program	Active duty Navy program patron				Non-active duty Navy program patron			
	$\bar{X}$	SD	Percent (%) reporting *agree*	Percent (%) reporting *disagree*	$\bar{X}$	SD	Percent (%) reporting *agree*	Percent (%) reporting *disagree*
CREDO Program	4.57	0.73	91.0	1.6	4.48	0.79	87.3	2.0
Recreation Program	4.63	0.61	95.2	0.8	4.60	0.72	92.8	1.7
Youth Programs	4.47	0.72	90.7	0.8	4.22	0.78	84.5	2.8
Child Development Program	4.35	0.90	88.6	6.1	4.25	0.90	82.7	3.8
Spouse Employment Assistance Program (SEAP)	4.34	0.86	82.8	3.4	4.25	0.83	83.6	2.4
NCP	4.27	0.84	84.8	3.2	3.90	0.88	64.4	3.1
Clinical Counseling Program	4.22	0.82	84.3	2.6	4.17	0.92	79.8	4.0
Relocation Assistance Program	4.21	0.74	85.0	1.5	4.33	0.65	91.9	0.9
Food and Hospitality Program	4.12	0.84	82.5	4.5	4.01	0.82	75.3	3.4
Financial Management Program	4.04	0.81	78.9	3.2	4.02	0.95	77.8	3.4
Nutrition Education and Galley Services Program	3.99	0.93	74.4	5.6	4.08	0.81	78.8	3.8
TAMP	3.86	0.89	69.6	6.0	3.96	0.95	71.5	6.1
Deployment Support Program	3.77	0.86	66.0	5.7	4.30	0.82	87.9	3.0

Note. Mean scores range from 1 = *strongly disagree* to 5 = *strongly agree*. Program totals are ordered highest to lowest by active duty Navy mean score.

Table 5
Program patron rating of *program effect on readiness* item

Program	Active duty Navy program patron				Non-active duty Navy program patron			
	$\bar{X}$	SD	Percent (%) reporting *agree*	Percent (%) reporting *disagree*	$\bar{X}$	SD	Percent (%) reporting *agree*	Percent (%) reporting *disagree*
CREDO Program	4.38	0.75	87.6	1.5	4.29	0.83	81.5	2.1
Youth Programs	4.08	0.85	55.5	11.1	3.78	1.09	81.1	5.5
Recreation Program	4.07	0.93	71.3	4.1	3.93	1.01	67.5	7.0
NCP	4.07	0.85	70.9	3.0	3.95	0.85	76.1	3.4
Spouse Employment Assistance Program (SEAP)	3.97	0.78	69.0	0.0	4.33	0.62	93.3	0.0
Child Development Program	3.93	0.96	76.1	9.1	4.03	0.68	79.3	0.0
Relocation Assistance Program	3.92	0.74	74.6	2.2	4.17	0.83	73.9	0.0
Nutrition Education and Galley Services Program	3.91	0.91	73.7	6.1	3.96	0.90	72.8	5.2
Clinical Counseling Program	3.89	0.88	68.1	5.4	3.75	1.11	66.7	8.3
Financial Management Program	3.87	0.83	73.0	5.3	3.95	0.96	73.3	5.5
TAMP	3.85	0.81	72.8	5.5	3.50	0.83	52.4	11.9
Deployment Support Program	3.82	0.78	70.6	4.1	4.00	0.82	75.0	0.0
Food and Hospitality Program	3.65	0.99	57.1	10.6	3.73	0.93	61.2	8.2

Note. Mean scores range from 1 = strongly disagree to 5 = strongly agree. Program totals are ordered highest to lowest by active duty Navy mean score.

Table 6
Correlation between program RFB mean score and readiness, QOL, and retention intent items (*$p < 0.05$, two-tailed. **$p < 0.01$, two-tailed)

Program	RFB $\bar{X}$	SD	Correlation to readiness (r)[a]	Correlation to QOL (r)[b]	Correlation to retention intent (r)[a]
Deployment Support Program	3.77	0.72	0.78**	0.77**	0.24*
Clinical Counseling Program	3.99	0.76	0.75**	0.68**	0.34**
Youth Programs	4.31	0.58	0.73**	0.72**	0.21**
Financial Management Program	3.86	0.70	0.72**	0.70**	0.35**
Recreation Program	4.23	0.64	0.71**	0.59**	0.27**
Nutrition Education and Galley Services Program	3.88	0.83	0.70**	0.72**	0.27**
Food and Hospitality Program	3.91	0.74	0.67**	0.70**	0.18**
Relocation Assistance Program	3.88	0.63	0.67**	0.68**	0.28**
NCP	4.10	0.69	0.66**	0.70**	0.24**
Spouse Employment Assistance Program (SEAP)	4.14	0.64	0.66**	0.54**	0.20
Child Development Program	4.22	0.78	0.65**	0.81**	0.26**
CREDO Program	4.42	0.63	0.65**	0.69**	0.34**
TAMP	3.50	0.83	0.63**	0.60**	0.21**

Note. Mean scores range from 1 = strongly disagree to 5 = strongly agree. Programs are ordered highest to lowest by program correlation to readiness.
[a] Correlation pertains to active duty Navy program patrons.
[b] Correlation pertains to all respondents including active duty Navy program patrons and non-active duty program patrons.

Financial Management, CREDO, and Clinical Counseling, while Food and Hospitality and TAMP showed the lowest (yet significant) correlations.

3. Discussion

Results from the present study support earlier military program evaluation research in demonstrating a positive relationship between QOL programs and its impact on Service members and their families. Specifically, this study shows the positive effect of Navy QOL programs on Sailors and their families. This study also successfully employed a cluster evaluation methodology and measures, modified from Kerce et al. (1999). For those RFB items that were used in both studies, results show that military recreation, child-care, youth, and employment assistance programs are rated most positively by program users. One program, CREDO, is rated more positively among Navy patrons than Kerce et al. (1999) report for USMC patrons.

CREDO, Recreation, SEAP, Child Development, Youth Programs, and NCP were consistently among the most positively rated programs whereas TAMP, Food and Hospitality, Nutrition Education and Galley Services, and Deployment Support were rated among the lowest. This general pattern of results is also seen in Kerce et al. (1999) in their work with USMC QOL programs. One reason for this may be that the program itself may facilitate a selection effect whereby the personal characteristics of individual program participants affect how an individual program is rated (Cook & Campbell, 1979). Patrons of programs that are associated with positive events (spiritual, personal, physical, or professional development) may be different in terms of affectivity (positive or negative) than users of programs that are associated with negative events (career transitions, deployment and family separation). This general

pattern of results may also be due to the nature of program services delivered. If one views the services of a program as a 'treatment' there may be varying levels of treatment delivery depending on the program (Rossi & Freeman, 1985). This study sought to control for this by asking program managers to administer surveys to intensive program users but there still may be differences in the program treatment and perceived benefits.

The relationship between RFB index mean scores for each program and the outcome items (QOL, readiness, and retention intent) shows the programs appear to be strongly related to readiness and QOL and modestly (yet significantly) related to retention intent. One program, SEAP, did not show a significant correlation to retention intent but due to the nature of the program patrons (predominantly non-military dependents), the number of program users who had a military retention decision to make was somewhat small ($n = 41$).

There are several limitations to this study that are worth noting. Since the project sponsor was primarily interested in the effect of these QOL programs on program users, there are no data that reflect the perceptions of those who chose not to use the programs. Navy MWR conducts an annual customer satisfaction survey and findings typically suggest that the reason Sailors do not use an MWR program is due to a lack of interest in the program or because they already use a similar program in the civilian community (Rosenfeld & Uriell, 2000).

The research design for this evaluation employed a post-test only using a purposive sampling approach (Levy & Lemeshow, 1999) where individuals were selected who were considered to be most representative of the population as a whole. We selected this approach for several reasons. First, more than half of the programs being evaluated do not know whom is going to be using their program at a given time—customers use the program on a walk-in basis.

Second, some programs serve a large number of customers (in excess of 200 per day) while other programs may serve very few customers (5–6 customers/day). Since the programs that the Navy has the least information on are the small programs that serve few customers, we decided to orient the sampling to ensure that we could collect enough data from the small programs to learn about the program's usefulness. Third, program managers wanted to ensure that those customers evaluating their program had enough knowledge about their program to represent a 'program user'. Although non-probability sampling limits the generalizability of results to the population of program users, it is recognized approach when faced with the design challenges listed above (Henry, 1990). Given the limitations associated with non-probability sampling, we were able to answer the primary research question—for those people who choose to use a QOL program, how well does it meet their needs and what impact does it have on higher-order organizational outcomes (i.e. overall QOL, readiness, and retention intent).

Sample sizes for some programs are somewhat low, such as Deployment Support and SEAP. The number of participants for Deployment Support was low for two reasons. First, we asked base-level program managers to administer the surveys to patrons/clients who are *experienced program users* (more than casual users). In the case of Deployment Support, this would have been a Sailor who not only attended a workshop but also may have requested individual counseling. This type of program then shifts from being a group-administered program to a one-on-one program (similar to Clinical Counseling). Second, data collection unknowingly occurred during a time in the Navy deployment cycle when few Deployment Support workshops were being administered at the two largest Navy bases in the world. SEAP had a low number of intensive program users for reasons much like Clinical Counseling and Deployment Support. Casual program patrons use the program in-group settings whereas intensive program users receive individualized attention.

There were also possibilities that there were inconsistencies in how the surveys were administered at the program sites. Although a member of the evaluation team conducted on-site survey administration training and left a standardized set of written instructions with program managers, variations in the administration of surveys may have occurred that could have affected participant selection and the range of sub-programs included in a program's evaluation. For example, the study methodology called for surveys to be distributed evenly across all sub-programs within a program. MWR Fitness and Recreation participants tended to be users of the Fitness Center (37%), the Bowling Center (21%), and the Recreation Center (10%) whereas users of Navy golf courses (2%), riding stables (0.6%), and intramural sports activities (0.2%) received very few surveys.

In spite of these limitations this cluster evaluation

approach serves as a major improvement to the way QOL program managers and resource sponsors previously collected program evaluation data. The main strength of this research study is the use of a common metric to evaluate dissimilar QOL programs. For the first time, Navy QOL program managers and resource sponsors can compare programs on the basis of how patrons rate them on meeting higher-order QOL objectives. With the addition of funding, staffing, and patronage data, program managers could determine which programs are the most high yield/cost effective programs. A challenge for the Navy, similar to what Kerce et al. (1999) found with the USMC, is that many QOL programs do not have a standardized way of counting program users and thus cannot calculate a cost per user index. Ongoing Navy QOL program development teams (termed *Integrated Process Teams*) are examining this issue and working toward a solution.

Another major advantage of this research approach centers on the survey instruments—brief surveys that are easy to administer and can be completed in 5–7 min. The surveys include only questions necessary to answer the overarching study questions—evaluating QOL programs and their impact on organizational outcomes.

4. Conclusions

This is the first time that the Navy has utilized a common set of measures to evaluate a broad range of programs. In the past, headquarters-level program managers have had to sort through findings from studies focused on individual program areas (e.g. MWR, FFSC, NCP, and CREDO conducted evaluations with measures that assessed their ability to provide recreation, family support, education, and personal/spiritual development, respectively). Few studies examined higher-order objectives (RFBs in this study) or specific organizational outcomes (QOL, readiness, and retention plans). This common set of measures allows for a multi-program evaluation that provides headquarters-level decision-makers information regarding those programs that are best meeting Sailor/family member needs. This is especially important given the Navy's support of the Revolution in Business Affairs and their new emphasis on standards, metrics, and measures of effectiveness (Froman, 2000).

These methods of program evaluation and subsequent results have proven useful in resource allocation for one program thus far. In late 1999 and early 2000, the CREDO program was facing a significant budget cut with more to come unless they provided data to prove their contention that CREDO was rated positively by patrons and significantly impacted Sailor readiness and/or retention intent. The Navy Chief of Chaplains office used preliminary results from the study to demonstrate the utility of the CREDO program to Navy organizational outcomes. As a result, funding for the CREDO program was increased and the

number of CREDO retreat locations increased from nine to 10 (J.W. Poole, personal communication, June 28 2000).

This research approach could be used in military and non-military settings where there is the need for a holistic, coordinated evaluation of programs with seemingly unique objectives. Specifically, other branches of the U.S. military forces (i.e. Army and Air Force) as well as foreign military forces could apply this evaluation approach when examining the impact of their inventory of QOL programs on their people. Additionally, private sector organizations that offer a broad array of employee support programs might consider such an evaluation approach.

Future research could examine the possible cumulative effect of a number of QOL programs on Sailor retention intent. Evaluated in a piece meal fashion, each program had a small but significant relationship with retention intent. It could be that one QOL program may not have a sizable effect on one Sailor but several programs used by one Sailor may have an impact—similar to a dosage effect. It is in this way that the effect of the full scope of Navy QOL programs can be thought of as one broad program—the Navy's QOL program.

Acknowledgments

References

Cavin, E. S. (1987). *A study of the marine corps family programs* (CNA Report 139). Alexandria, VA: Center for Naval Analysis.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin Company.

Froman, V. Z. (2000). *Shore installation management: A new way ahead.*

Briefing presented to Navy Ashore Installation Management Directors, Washington, DC.

Henry, G. T. (1990). *Practical sampling*. Newbury Park, CA: Sage Publications.

Kerce, E. W. (1998). *Assessment of USMC quality of life (QOL) program contributions to readiness, performance, and retention: Volume I, design and methodology* (TN-98-6). San Diego, CA: Navy Personnel Research and Development Center.

Kerce, E. W., Sheposh, J., & Knapp, S. (1999). *Assessment of USMC quality of life (QOL) program contributions to readiness, performance, and retention: Volume II, pilot test results* (TN-99-5). San Diego, CA: Navy Personnel Research and Development Center.

Levy, P. S., & Lemeshow, S. (1999). *Sampling of populations: Methods and applications* (3rd ed). New York: Wiley.

Reed, C. S., & Brown, R. E. (2001). Outcome-asset impact model: Linking outcomes and assets. *Evaluation and Program Planning, 24*, 287–295.

Rosenfeld, P., & Uriell, Z. A. (2000). *Results of the 2000 Navy Morale, Welfare and Recreation Customer Survey*. Briefing presented to the Assistant Commander, Navy Personnel Command for Personnel Readiness and Community Support, Millington, TN.

Rossi, P. H., & Freeman, H. E. (1985). *Evaluation: A systematic approach* (4th ed). Newbury Park, CA: Sage Publications.

Sanders, J. R. (1997). Cluster evaluation. In E. Chelimsky, & W. R. Shadish (Eds.), *Evaluation for the 21st century: A handbook* (pp. 396–404). Thousand Oaks, CA: Sage Publications.

Schwerin, M. J., Michael, P. G., Glaser, D. N., & Uriell, Z. A. (2002a). *Quality of life (QOL) program contributions to readiness and retention. Volume II: Summary report*. Navy Personnel Research, Studies, and Technology (PERS-1) Technical Report (in press).

Schwerin, M. J., Michael, P. G., & Uriell, Z. A. (2002b). *Navy quality of life (QOL) program contributions to readiness and retention. Volume 1: Design and methodology*. Navy Personnel Research, Studies, and Technology (PERS-1) Technical Report (in press).

Sharma, R. (1994). The navy retention/separation survey: Enlisted responses for FY 1990 through FY 1992 (CNA research memorandum 94-28). Alexandria, VA: Center for Naval Analysis.

Further Reading

Michael, P. G., Schwerin, M. J., & Uriell, Z. A. (2001). *Nutrition, education & galley program assessment: Preliminary report*. Navy Personnel Research, Studies, & Technology (PERS-1) Technical Note. In press.

Schwerin, M. J. (2002). *The impact of Morale, Welfare, & Recreation (MWR) programs on aviator retention*. Navy Personnel Research, Studies, & Technology (PERS-1) Technical Report. In press.

Schwerin, M. J., Michael, P. G., & Uriell, Z. A. (2001). *Navy quality of life (QOL) program contributions to readiness and retention. Volume I: Design & methodology*. Navy Personnel Research, Studies, & Technology (PERS-1) Technical Report. In press.

Schwerin, M. J., Michael, P. G., Glaser, D. N., & Uriell, Z. A. (2001). *Navy quality of life (QOL) program contributions to readiness and career intentions. Volume II: Pilot test results*. Navy Personnel Research, Studies, & Technology (PERS-1) Technical Report. In press.

Wilcove, G. L., Wolosin, D. G., & Schwerin, M. J. (2001). *Development of a new model for quality of life using structural equation modeling*. Navy Personnel Research, Studies, & Technology (PERS-1) Technical Report. In press.