# Affective Factors in Generation of Tailored Genomic Information

Nancy Green
University of North Carolina at Greensboro
Dept. of Mathematical Sciences
Greensboro, NC 27402-6170 USA
336-334-5836
nlgreen at uncg.edu

## ABSTRACT
GenIE, a prototype intelligent system currently under development, is intended to assist genetic counselors by creating editable first drafts of tailored patient letters. Informed by individualized user models, GenIE will employ natural language text generation techniques. This paper discusses the affective component of the user models, and its relationship to the normative component. The normative component contains biomedical information, while the affective component describes emotion, coping behavior, risk perception, and other affect-related variables.

## 1. INTRODUCTION
The role of genetic counselors in the USA is to provide their clients with nondirective, tailored information on topics such as genetic testing, diagnosis and treatment of genetic disorders, and inheritance risks. According to Kessler [14], this kind of information may trigger a variety of negative emotions in clients, including low self-esteem, feeling of loss or sadness, blame, anger, fear, shame, worry, and guilt; these emotions in turn may trigger various coping strategies. Affective factors may influence a client's ability to comprehend information, perception of risk, and decision making. Thus, it is important for genetic counselors to present information "empathically" (p. 49).

In current practice, after meeting with the client the counselor writes a one- to two-page tailored print document addressed to the client, the *patient letter*, summarizing information provided and clinical services performed [1]. GenIE, a prototype intelligent system currently under development, is intended to assist genetic counselors by creating editable first drafts of patient letters. Employing natural language text generation techniques [4,13], and informed by an individualized user model, GenIE will synthesize a tailored letter. Here *user* refers to the primary addressee(s) of a letter, i.e., the client(s), rather than the genetic counselor. Clients may be patients or parents/guardians of patients; secondary readers of a letter may include other family members and those who interact with the patient such as caregivers and teachers.

GenIE's graphical user interface will be designed to enable the counselor to quickly provide two types of client-specific information. Biomedical information about the patient and the patient's biological family, such as symptoms, test results, and diagnosis, will be used by the system to tailor the *normative* component of the *user model* (NUM) for a particular client [10]. Other observations about the client's situation will be used to tailor the *affective* component of the *user model* (AUM). The motivation for incorporating affect into the user model is to enable GenIE, when drafting patient letters, to express the counselor's empathy for the client.

This paper presents some requirements and a preliminary design for the AUM. First we describe the NUM and how we plan to augment it with the AUM. Next, drawing upon research on psychological issues in genetic counseling, we outline types of variables to be included in the AUM. Then we present the intended role of the AUM in natural language generation of patient letters.

## 2. ADDING AFFECTIVE FACTORS TO A NORMATIVE USER MODEL
In GenIE, a NUM will be created by updating a copy of the system's generic genomic health information knowledge base with client-specific biomedical information [10]. The design of the generic knowledge base was informed by analysis of a corpus of genetic counseling patient letters. The letters were encoded using an annotation scheme we developed for representing the biomedical content presented in genetic counseling patient letters [8]. Biomedical concepts are represented as probabilistic belief network variables [15]. Types of event/state variables include *genotype*, representing the number (0, 1, or 2) of mutated alleles (copies) of a particular gene, *symptom*, *history*, test *result*, etc. A prototype knowledge base is being implemented as a qualitative probabilistic network (or QPN) [20,6]. A QPN uses qualitative constraints instead of numeric probability distributions. (Qualitative constraints are easier to obtain and are sufficient for our user model.)

For illustration, let us consider a case study about the R. family presented in [19]: the R's youngest child was diagnosed as having Phenylketonuria (PKU) at the age of six months. An in-born error of metabolism, PKU is transmitted by autosomal recessive inheritance. (An autosomal recessive disorder is caused by having two mutated alleles of a gene. Everyone inherits one allele from the mother and the other from the father. If a parent has only one mutated allele, then he or she may transmit the mutation to the child but will not appear to be affected by the disorder.) PKU can be detected by testing at birth or later. The longer the disorder remains undetected, and thus untreated, the more likely it is to have serious, irreversible effects such as chronic illness and mental retardation. Treatment consists of adherence to a highly restricted, low protein diet during childhood. Since the R.'s child was not diagnosed until the age of six months, he already has been affected with mild mental retardation.

Figure 1 depicts a NUM for PKU. White ovals represent event/state variables and rectangles represent decision variables [12]. Variable types are shown in bold font. Family member terms (e.g., *Mother*) are attached to variables to indicate which member of the client's family is described. Arcs leading to event/state variables represent qualitative influence; arcs leading to decision variables represent information used in decision making. As shown, the genotypes of the mother and father influences the genotype of their child, which influences the occurrence of symptoms of PKU such as mental retardation. Evidence that the child has PKU may be provided by test results. Knowledge of test results is used in deciding to provide dietary treatment, which may prevent further mental retardation. Unfortunately, adherence to the strict diet may influence quality of life (QOL) in terms of parental stress and behavior problems in the child.

So far, we have described only generic aspects of PKU. A NUM tailored for the R. family's case could be created when the counselor enters the following information through GenIE's graphical user interface: the patient was observed to have mild mental retardation, the patient's PKU test result is positive, the patient is believed to have PKU (i.e., two copies of the mutation), and, since they do not appear to have PKU, the patient's parents are believed to be PKU-carriers (i.e., to have only one copy of the mutation each). The values of variables representing this tailored NUM are shown in parentheses in the figure. Note that not all values need be supplied by the counselor. For example, the potential effects of treatment on QOL in terms of parental stress and behavior problems would be predicted as likely from the other information.

Future effects of PKU on the R. family's child's health could be mitigated by continued adherence to dietary treatment. However, adherence often depends upon a family's understanding of the role of diet in management of PKU; also, psychological issues may interfere with continued adherence [19]. The case study of the R. family illustrates a variety of possible psychological effects on the parents of a child diagnosed with PKU: self-blame about transmitting the mutation, guilt induced by the child's behavior problems in reaction to dietary restrictions, feelings of stigmatization about the child's mental retardation, worry about the child's future, denial of the child's mental retardation, anger projected onto physicians, and questioning the value of continuing dietary treatment.

We contend that a system's ability to draft patient letters conveying the counselor's empathy depends upon having access to this sort of information about a client. Our goal is to design an affective component of the user model (AUM) that will be integrated with the NUM. Based on review of the literature on psychological aspects of genetic counseling and on risk perception, we plan to implement stereotypical affective user models for different genetic disorders. The stereotypical AUM then could be tailored for a particular client using information supplied by the counselor through GenIE's graphical user interface.

For a preview of what we mean see Figure 2, which adds an affective component to the NUM shown in Figure 1. In Figure 2, gray nodes represent AUM variables, and the variables' values (shown in parentheses) represent the R. family's case. Values not observed by the counselor could be inferred based upon qualitative constraints. For example, if the counselor observed a coping strategy of denial, the system could infer that it is likely that there had been a negative parental response to the patient's mental retardation. Also, if behavior problems have been reported, then it could be predicted that the R. parents possibly will feel guilt about treatment; and that the guilt and resulting denial may result in the future in a decision to discontinue their child's dietary treatment. Thus, a letter tailored for the R. family might include strategies appropriate to the R's response to their child's mental retardation or tips on dealing with typical behavior problems resulting from the diet.

Note that it would not be sufficient to use only a stereotypical AUM for each disorder such as PKU. To illustrate why not, consider another case reported in [19]: the M. family's child was not diagnosed until the age of 23 months and as a result is moderately retarded. Due to the diagnosis, his younger sibling was also tested at that time and was also found to have PKU. However, with help from a PKU support group, the parents have accepted their childrens' condition. Thus, a letter tailored for the M. family should reflect these differences from the R. family.

## 3. TYPES OF VARIABLE IN THE AUM

In this section, we describe and motivate some variable types to be used in the AUM. *Emotion* type variables will be used to represent a client's emotion towards himself, another individual, or an event. Many descriptive emotion terms are used in the psychology literature; we do not know yet what level of specificity will be sufficient for our purposes. Emotion events may be triggered by an individual's appraisal of an event represented in the NUM. For example, "To be responsible for an innocent child's illness … may generate enormous guilt in the parents. This is particularly true if the disease is not transmitted by an autosomal recessive mode of inheritance and thus culpability can be ascribed to one rather than both of the parents" ([14], p. 24).

Emotions may lead to coping, which in turn may lead to other emotions. For example, "Parents who comply with the dietary and medical regimen [for PKU] are simultaneously reinforcing their sense of being responsible for the child's 'defect.' This tends to lead to avoidance behavior and a laxity with respect to adherence to the medical regimen which, in turn, leads to an increasing sense of guilt" ([14], p. 8). The *coping strategy* variable type will be used in the AUM to represent possible coping behaviors, such as those described in [3]. Coping strategies may be influenced by personality in combination with emotion events. The *personality* variable type will be used to represent personality types, according to a classification system relevant to genetic counseling.

Also, various *interpersonal* variables will be represented in the AUM. According to [14], "…genetic disorders have a major impact on family functioning … To a large extent, the interpersonal responses of parents are contingent upon their individual personalities, past responses to stressful events, the stage of the life cycle of the family, the presence (or absence) of a supportive social system and other important variables" (p. 10). For example, in the M. family case study discussed earlier, membership in the PKU support group could be represented as an interpersonal variable contributing to the parents' coping strategy of acceptance of their childrens' condition.

In some cases, a client's prior knowledge may need to be represented in an AUM. For example, an individual who knows that Huntington's disease (HD) is terminal and incurable may have a different reaction to a diagnosis of HD than someone who is not aware of that. In the former case, "The diagnosis of a genetic disorder may seem like a sentence of doom … There is a sense of fatalism, a hopelessness about the future" ([14], p. 24).

In addition to general factual knowledge, an individual may have personal experience that contributes to his appraisal of a diagnosis. For example, "… a crucial variable in the at-risk [of HD] individual's adjustment to their risk is the nature of the earlier exposure to the disorder in their childhood. When the affected parent was able to remain functional in the home and/or the quality of care that was perceived to have been received was good, the later reactions of the at-risk person tended to be relatively optimistic" ([14], p. 9). We propose to represent factual knowledge and personal experience using a *belief* variable type in the AUM.

Actual risk is represented in the NUM. However, perceived risk may differ from actual risk. For example, one study found that women who had a first-degree relative with breast cancer overestimated their risk before and after counseling [16]. In order to model discrepancies between actual risks, represented in the NUM, and the client's perception of risks, the AUM will include *risk perception* variables. Risk perception may be influenced by psychological factors such as worry [11]. Unrealistically high perceived risk can interfere with adherence with recommended breast cancer prevention and surveillance strategies [16].

We conclude this section with two more examples of user models. Figure 3 illustrates an integrated stereotypical NUM-AUM for Huntington's Disease (HD) based on information from [21]. (Unlike PKU, HD is transmitted by autosomal dominant inheritance, i.e., inheritance of one copy of the HD mutation is sufficient to cause the disease. To simplify the diagram, nodes representing the patient's ancestors with HD have been collapsed into the node labeled *Family history of HD*.)

As shown in Figure 3, an individual's decision to be tested may be influenced, not only by his perceived risk of having inherited the mutation, but also his knowledge that HD is fatal and his risk perception of a short life expectancy and low quality of life. Thus, an at-risk individual may prefer not to know if he has inherited the HD mutation and may decide not to be tested. In addition, the model shows that after testing, knowledge of test results and the above affective factors may contribute to a strong, negative emotional response; for some personality types, that may lead to suicidal thoughts. (According to [21], at-risk persons who had personal experience of a parent ill with HD feared having the symptoms of HD more than death.) The model also describes another possible coping strategy of those who perceive themselves to be high risk: intense productivity to compensate for a shorter life expectancy.

Figure 4 illustrates part of a stereotypical NUM-AUM for breast cancer (BC) due to an inherited mutation of the BRCA1 or BRCA2 genes. Certain ethnic groups are at higher risk of carrying the mutation. Unlike the genetic disorders illustrated above, the predisposition to develop BC is inherited but not the disease itself. However, risk increases with age at a higher rate among those with the inherited mutation compared to those without it. Figure 4 models how perceived risk of developing BC in the future may be influenced by affective factors such as personal experience with a family member who had BC. Perceived risk may positively or negatively influence a patient's decision to have BRCA1/2 testing. In addition, Figure 4 represents potential costs of BRCA1/2 testing, including risk of discrimination or loss of insurance. As shown in the figure, a patient's perception of this risk also may influence the decision to have testing.

## 4. USE OF AFFECTIVE FACTORS

The motivation for adding an affective component to the user model is to enable GenIE, when drafting patient letters, to express the counselor's empathy with the client. A tone of empathy can be evoked by a text through its selection of content as well as its choice of stylistic options (words and phrasing). During text generation, the selection of content and stylistic options will be constrained by the user model.

We have identified several affect-related strategies for content selection in a corpus of patient letters written by genetic counselors that could be adopted in GenIE [9]. For example, one letter written to the parents of a patient points out that they had no control over which copy of the gene causing their child's disorder was inherited from each parent. Presumably, this was included in case the parents blamed themselves for their child's genetic disorder.

Baker et al. [1] present stylistic techniques for human authors that can be used when writing genetic counseling patient letters. Distance between the staff and patient can be reduced by use of personal expressions of sympathy instead of formulaic expressions (e.g., "We were so sad to learn of" instead of "We want to extend our condolences on"), or by reference to the staff (e.g., "We are sorry that your partner was unable to attend the visit and hope this letter will serve to review for him the information that was discussed" versus "This letter will also serve to review for your partner the significant information discussed") (p. 408). Distance can be reduced by referring to the addressee in the second person (e.g., "You and your children") and by use of first person instead of third person (e.g., "We referred you to …" instead of "A referral to … is indicated") (p. 411-2).

Baker et al. suggest conveying a greater distance between the patient and negative information in the letter. One technique is to emphasize the relationship of the diagnosis to the symptoms instead of to the patient (e.g., "Based on the physical features that were noted during the examination, you were diagnosed with Marfan syndrome" versus "You were diagnosed with Marfan syndrome") (p. 409). Another technique is to exclude direct mention of the patient from negative information (e.g., "Children with this condition tend to lose their hearing" instead of "Nisha is likely to lose her hearing") (p. 410). A third technique for increasing distance between the patient and negative information is use of passive voice (e.g., "Nystagmus was noted to be more pronounced during Christina's exam" versus "Christina is experiencing increased nystagmus…") (p. 411).

## 5. RELATED WORK

Genetic counseling patient letters are supposed to provide *nondirective* information, and so our system is not intended to

persuade the user to act in a certain way as in systems for generating arguments to promote healthy behavior, e.g., [7,18]. Our work has been influenced by recent research on computational models of emotion. Based on psychological theories of appraisal and coping, EMA is a computational model used to generate realistic behavior in human characters for virtual training environments [17]. However, EMA is not implemented as a probabilistic belief model, as is the user model in GenIE. A dynamic probabilistic model to monitor a user's emotions and engagement while playing an educational game is described in [5]. Based on a cognitive theory of emotion, the model enables changes in the user's emotional state over time to be represented. This model is used to select pedagogical interventions. Another dynamic probabilistic model of emotion, the model of [2] incorporates a finer-grained representation of emotion and is used to control the facial expressions of an embodied agent.

## 6. SUMMARY

GenIE, a prototype intelligent system currently under development, is intended to assist genetic counselors by creating editable first drafts of patient letters. Employing natural language text generation techniques, and informed by an individualized user model, GenIE will synthesize a tailored letter. We presented a preliminary design of how we plan to incorporate affective factors into the current design of the normative user model for GenIE.

## 7. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under CAREER Award No. 0132821.

## 8. REFERENCES

[1] Baker, D.L., Eash, T., Schuette, J.L., and Uhlmann, W.R. Guidelines for writing letters to patients. *Journal of Genetic Counseling*, 11(5), 2002, 399-418.

[2] Carofiglio, V., de Rosis, F., and Grassano, R. Dynamic models of mixed emotion activation. In D. Canamero and R Aylett (Eds), *Animating expressive characters for social interactions*. John Benjamins, in press.

[3] Carver, C.S., Scheier, M.F., and Weintraub, J.K. Assessing coping strategies: a theoretically based approach. *Journal of Personality Psychology*, Vol. 56, 1989, 267-283.

[4] Cawsey, A.J., Webber, B.L., and Jones, R.B. Natural language generation in health care. *J. Am Med Inform Assoc*, 4(6), 1997, 473-82.

[5] Conati, C. Probabilistic Assessment of User's Emotions in Educational Games. *Applied Artificial Intelligence*, 2002.

[6] Druzdzel, M..J., and Henrion, M. Efficient Reasoning in Qualitative Probabilistic Networks. In *Proc. 11th Nat. Conf. on AI*, 1993, 528-553.

[7] Grasso, F., Cawsey, A., and Jones, R.. Dialectical Argumentation to Solve Conflicts in Advice Giving: A Case Study in the Promotion of Healthy Nutrition. *Int .J. Human-Computer Studies* 53, 2000, 1077-1115.

[8] Green, N. A Bayesian network coding scheme for annotating biomedical information presented to genetic counseling clients. *Journal of Biomedical Informatics* 38, 2005, 130-144.

[9] Green, N. Argumentation in a Causal Probabilistic Humanistic Domain: Genetic Counseling. *Int. Journal of Intelligent Systems*. To appear.

[10] Green, N., Britt, T., Jirak, K., Waizenegger, D., Xin, X. User Modeling for Tailored Genomic e-Health Information. In *User Modeling 2006, Workshop on Personalisation for eHealth*. To appear.

[11] Gurmankin, A.D., Domchek, S., Stopfer, J., Fels, C., and Armstrong, K. Patients' Resistance to Risk Information in Genetic Counseling for BRCA1/2. *Arch Intern Med*, vol. 165, 2005, 523-529.

[12] Henrion, M., Breese, J.S., and Horvitz, J. Decision Analysis and Expert Systems. *AI Magazine*, Winter 1991, v. 12, no. 4, 64-91.

[13] Hüske-Kraus, D. Text Generation in Clinical Medicine – A Review. *Methods Inf Med* 42, 2003, 51-60.

[14] Kessler, S. (ed.) *Genetic Counseling: Psychological Dimensions*. Academic Press, NY, 1979.

[15] Korb, K., and Nicholson, A.E. *Bayesian artificial intelligence*. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[16] Lerman, C., Lustbader, E., Rimer, B., Daly, M., Miller, S., Sands, C., and Balshem, A. Effects of Individualized Breast Cancer Risk Counseling: A Randomized Trial. *Journal of the National Cancer Institute*, Vol. 87, No. 4, 1995, 286-292.

[17] Marsella S, and Gratch J. Modeling Coping Behavior in Virtual Humans: Don't Worry, Be Happy. In *Proceedings of Autonomous Agents for Multiagent Systems (AAMAS'03)*, 2003.

[18] Reiter, E. Robertson, R., and Osman, L. Lessons from a Failure: Generating Tailored Smoking Cessation Letters. *Artificial Intelligence* 144, 2003, 41-58.

[19] Schild, S. Psychological Issues in Genetic Counseling of Phenylketonuria. In Kessler (ed.), ch. 8.

[20] Wellman, M.P. Fundamental concepts of qualitative probabilistic networks. *Artificial Intelligence*, 44(3), 1990, 257-303.

[21] Wexler, N.S. Genetic "Russian Roulette": The Experience of Being "At Risk" for Huntington's Disease. In Kessler (ed.), ch. 12.

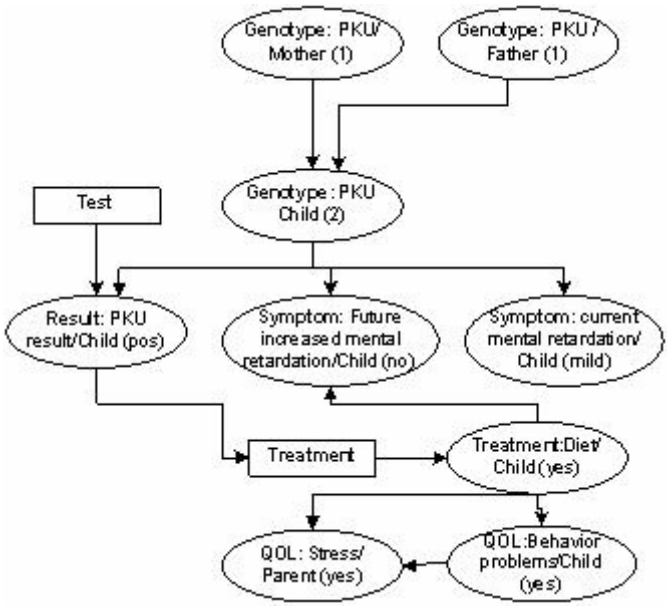


**Figure 1. Normative component of user model (NUM) for family R., a family with a child who has PKU.**

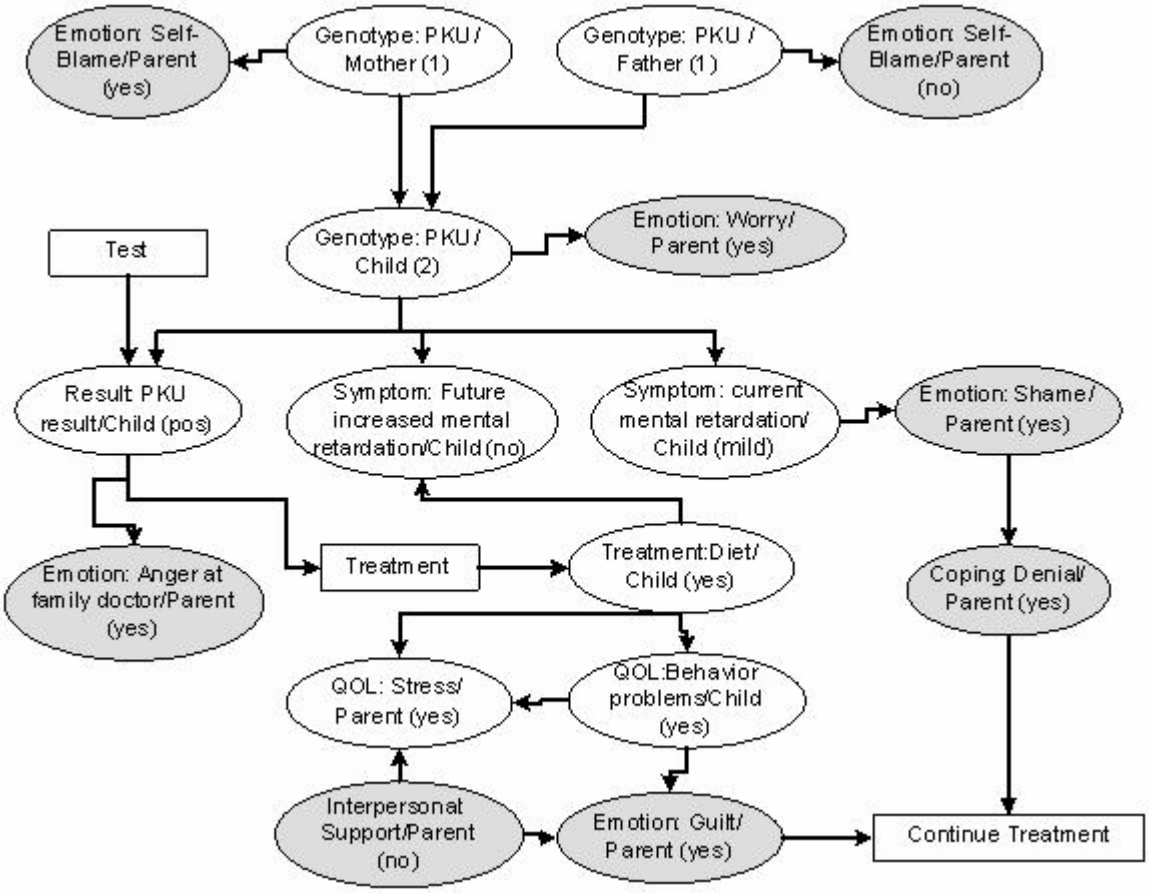


**Figure 2. Integrated normative and affective components of user model (NUM-AUM) for family R.**

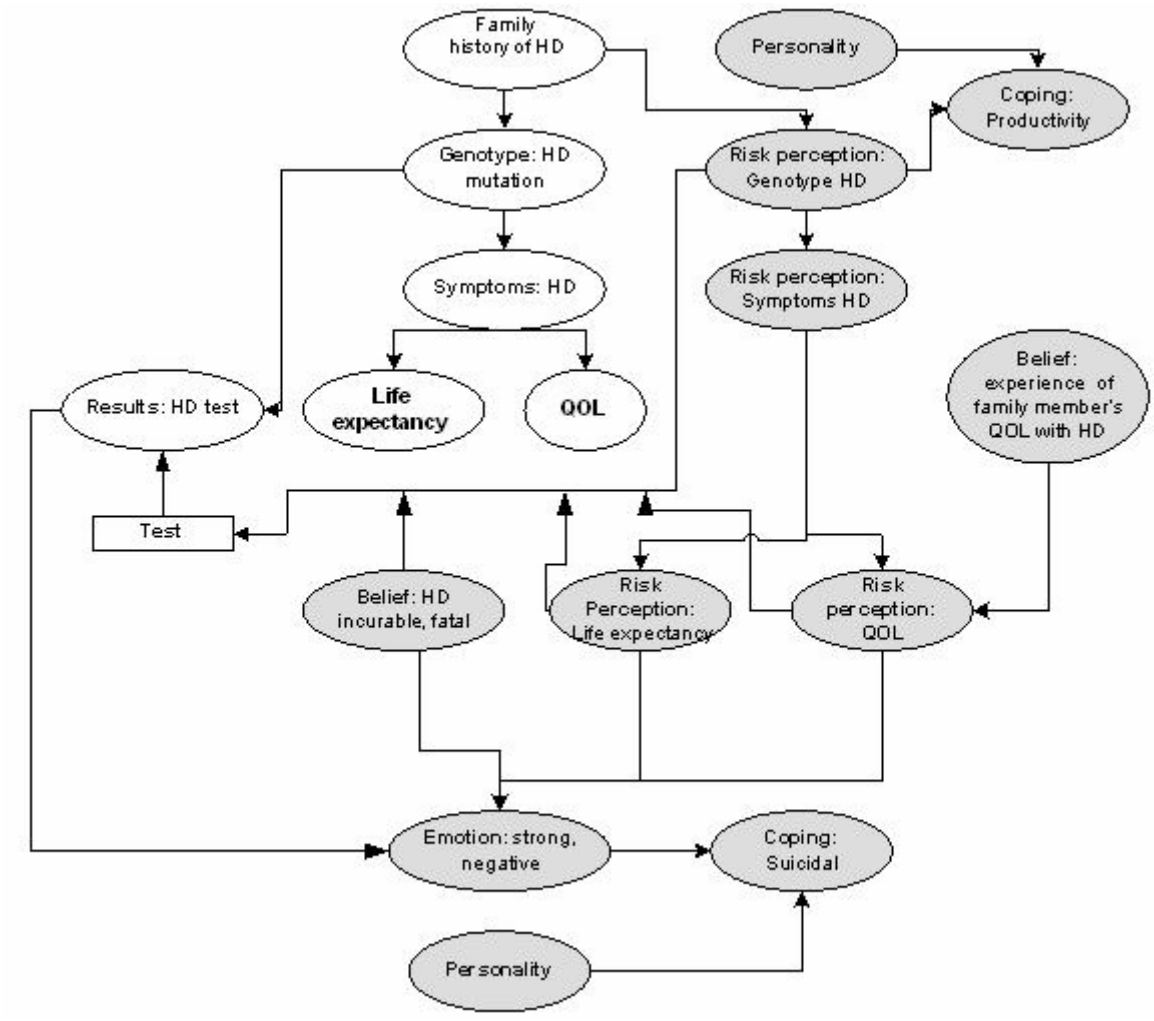


**Figure 3. Integrated normative and affective components of stereotypical user model (NUM-AUM) for HD.**

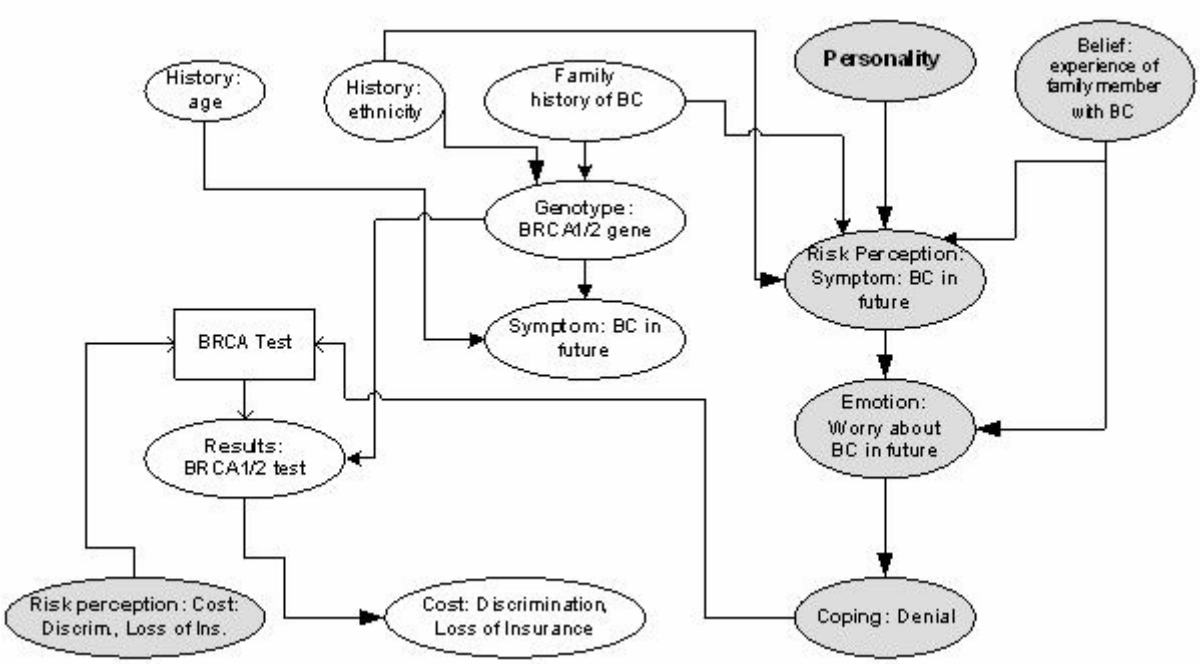


**Figure 4. Part of integrated normative and affective components of stereotypical user model (NUM-AUM) for BC.**